

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:41:00 ; Search time 55.96 seconds

(without alignments)  
1101.608 Million cell updates/sec

Title: US-09-913-414-4

Perfect score: 2807  
Sequence: 1 MFRSGAKPTFDQAVLVAR.....MNINGVISDPCGRDCRCPWS 555

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A.Geneseq\_032802.\*  
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2807	100.0	555	21	AA1982786 Caulobacter cresce
2	1195	42.6	580	20	AA1955922 Pseudomonas fluores
3	1193.5	42.5	583	20	AA1955819 Pseudomonas fluores
4	1178.5	42.0	588	17	AA1905297 Esterase secretory
5	572.5	20.4	710	20	AA1927215 Amino acid sequenc
6	569.5	20.3	707	18	AA1922153 ApxB protein. Ac
7	569.5	20.3	707	21	AA1951407 A. pleuropneumonia
8	567.5	20.2	710	20	AA1927214 Amino acid sequenc
9	555.5	19.8	560	22	AA1936153 Klebsiella pneumon
10	552.5	19.7	711	18	AA1922160 ApxB protein. Ac
11	552.5	19.7	711	21	AA1951413 A. pleuropneumonia

12	549.5	19.6	723	22	AA1933597 Pseudomonas aerugi
13	521.5	18.6	718	22	AA1938443 Salmoneella typhi c
14	457	16.3	571	22	AA1939313 Human polypeptide
15	454	16.2	578	22	AA1936908 Staphylococcus aur
16	453	16.1	578	22	AA1936908 Human ABC transpor
17	450.5	16.0	574	21	AA1954454 Amino acid sequenc
18	450	16.0	843	22	AA1900011 Human ABCB12 trans
19	449	16.0	843	22	AA1900010 Human ABCB12 trans
20	443	15.8	665	22	AA1952752 Escherichia coli p
21	440	15.7	582	22	AA1934522 E. coli cellular p
22	435	15.5	582	22	AA1934513 S. epidermidis ope
23	422.5	15.1	548	22	AA1931616 Haemophilus influe
24	419.5	14.9	587	22	AA1935367 Haemophilus influe
25	419.5	14.9	587	22	AA1935367 Haemophilus influe
26	418.5	14.9	766	22	AA192437 Human ATP binding
27	418.5	14.9	766	22	AA192437 Human ATP binding
28	417.5	14.9	591	22	AA1941099 Human polypeptide
29	415.5	14.8	766	22	AA197163 Amino acid sequenc
30	415.5	14.8	766	22	AA197163 Amino acid sequenc
31	414	14.7	1034	20	AA193792 S. aureus polypept
32	413.5	14.7	766	22	AA192442 Human ATP binding
33	409	14.6	621	22	AA192266 C. glutamicum prote
34	409	14.6	621	22	AA192266 Corynebacterium gl
35	407	14.5	1247	22	AA1969097 Drosophila melanog
36	404.5	14.4	866	22	AA1971210 Pseudomonas aerugi
37	403.5	14.4	603	22	AA193706 Micromonospora eve
38	403.5	14.4	613	22	AA1904887 Human protein sequ
39	403.5	14.4	722	22	AA192941 Human membrane tra
40	403	14.4	718	21	AA1971059 Human protein sequ
41	402	14.3	735	22	AA193546 Human ATP-binding
42	401.5	14.3	747	20	AA1959585 Human ATP-binding
43	401.5	14.3	747	21	AA193355 Leucocin A ATP tra
44	401	14.3	717	19	AA1959008 ABC transporter (1
45	401	14.3	717	20	AA194886

#### ALIGNMENTS

RESULT	1	ALIGNMENTS
AA192786	AA192786 standard; protein; 555 AA.	
XX	AA192786;	
XX	22-DEC-2000 (first entry)	
XX	Caulobacter crescentus ABC transporter protein, Rsad.	
XX	ABC transporter protein; Rsad; ATP binding; type I secretion system;	
XX	C-terminal secretion signal; Caulobacter host cell; biofilm formation;	
XX	bioreactor; recombinant protein expression host; sewage treatment;	
XX	wood pulping; fish vaccine.	
XX	Caulobacter crescentus.	
XX	WO200049163-A1.	
XX	24-AUG-2000.	
XX	21-FEB-2000; 2000WO-CA00173.	
XX	19-FEB-1999; 99CA-2261186.	
XX	(UYBR-) UNIV BRITISH COLUMBIA.	
XX	Smit J;	
XX	WPI; 2000-571987/53.	
XX	Caulobacter host cell for expression and secretion of a heterologous	
XX	polypeptide, useful for treating sewage, waste water and in the pulping	
XX	process -	

XX Claim 1; Page 35; 46pp; English.

CC The invention relates to a Caulobacter host cell for expression and  
PS secretion of a heterologous polypeptide. The Caulobacter host cell  
XX comprises at least one surface layer transport protein having an  
CC amino acid sequence homologous to the Caulobacter crescentus Rsad or  
CC RsaE proteins (AAB22786, AAB22787). The host cell further comprises a  
CC DNA construct comprising DNA encoding a polypeptide heterologous to a  
CC surface layer (S-layer) protein of the cell, from, and operably linked  
CC to, a DNA encoding a Caulobacter surface layer protein secretion signal,  
CC with the proviso that when the cell comprises transport proteins having  
CC the same sequence as both the Rsad and RsaE proteins, the secretion  
CC signal is not from the Caulobacter crescentus surface layer protein  
CC (RsaE). The invention also encompasses a method for identifying a  
CC Caulobacter suitable for use as a host cell for expression and secretion  
CC of a heterologous polypeptide comprising the detection of homologues of  
CC the rsad or rsae genes using a rsad/rsae hybridisation probe. The host  
CC cell is used for the expression and secretion of a heterologous  
CC polypeptide. The modified Caulobacter cells may be used to treat sewage  
CC and waste water. The cells may also be grown in wood pulp suspensions and  
CC then used in the wood pulping process. They can also be used in fish  
CC vaccines. Caulobacter species are able to form biofilms, and can attach  
CC themselves to surfaces without producing the extracellular enzymes or  
CC polysaccharide slimes that are characteristic of most other  
CC surface-attached bacteria. Caulobacter are therefore particularly suited  
CC for use in bioreactor systems. The present sequence represents the  
CC Caulobacter crescentus ABC transporter protein, Rsad. This protein is  
CC embedded in the inner membrane of Caulobacter crescentus and contains an  
CC ATP-binding region. It recognises the C-terminal secretion signal  
CC (AAB22785) of the substrate protein (e.g., Rsad), and hydrolyses ATP  
CC during the transport process.

XX Sequence 555 AA:

Query Match 100.0%; Score 2807; DB 21; Length 555;  
Best Local Similarity 100.0%; Pred. No. 8.9e-232;  
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKRSQAKPTFDDAVICVAPAVITAMVSPFINILAVSPLYMLOYVDRVLTSMNSTL 60  
DB 1 mfkrsqakptfddavicvavpavitamvspfinilavsplymloyvdrvtstnsvcl 60  
QY 61 IVLVYICVLEFLVYGLLEALRTQVLYVRGSLKFDGVARDPFKSVYDSTLSRKGIQGA 120  
DB 61 IvlvYicvLfLygllEalrtqVlyvRgSlKfdgvarDpFKsvYdStLSrkGIqgAfr 120  
QY 121 DMDQVREMTGGLAFCDAPMTFVYVSMVLPFFGLAIIICIIIFGLAVMMDNAKN 180  
DB 121 dmdqvremtgglafcdapmtfvYvsmVlpffGLaIIICIIIFglAVmmdnAKn 180  
QY 181 PLOMATIASAONDAGSTLIRNAEVMKAMWGGLQARWRARRDROVMOAASDAGAV 240  
DB 181 plomatiasaondagstlIRnaeVmKamWgGLQarWRarRDroVmoAasdaGav 240  
QY 241 MSGIKVFENIVOTLLIGGAVLADGKISAGAMTAGSILVGRALAPIEGAVGQMKNTYGA 300  
DB 241 msgikvfEnivotLLiGGavLadGkISaGAMtAGsILvGRaLAPIeGavGQmKNtYga 300  
QY 301 RGAMRLOTMLREKSSADNDHPLPEPRGVLSAEASILPPGACQPTMQASFRIDAGAAV 360  
DB 301 rGAMrLoTmLrEkSSADnDhPlPePrGvLsAEaSiLpPgAcQpTmQaSFriDaGAAv 360  
QY 361 ALVGSASAGKSLIRGIVGWPCAGVIRLDGYDKOMDPKLRHVGVLPODIELFSGT 420  
DB 361 alvgsasagKslIRgIvGwPcAGvIRldGyDKoMDpKLRhVgVLpOdIElFsgT 420  
QY 421 VAONITARTEFESQEVIAEATLAGVHEMIOSLPMGYDPAIGEGASLSGGORLALARA 480  
DB 421 vaonItARtEfESqEvIAeATlAGvHEmIOsLpMGyDPAIGeGASlSGGORlAlaRa 480  
QY 481 VERMPALLVDEPNASLDQVGEVALMEAMKRLKAARKRTIVATRHVNLDAQDYLIMVING 540

DB 481 vfmppallvdepnasldqvgEvalmeamkrlkaarkrtivatrHvnlDaqDylImvIng 540  
QY 541 GVISDFGERDRCWPS 555  
DB 541 gvisdFgerdrcwps 555

RESULT 2  
ID AAY55922 standard; protein; 580 AA.  
XX AAY55922;  
AC AAY55922;  
DT 15-FEB-2000 (first entry)  
XX  
DE Pseudomonas fluorescens ABC transporter cassette II component A.  
XX  
KW ATP-binding cassette; transporter; operon; lipBCD; Serratia marcescens;  
KM microbe; protein secretion.  
XX  
OS Pseudomonas fluorescens.  
XX  
PN JP11276172-A.  
XX  
PD 12-OCT-1999.  
XX  
PF 27-MAR-1998; 98JP-0080597.  
XX  
PR 27-MAR-1998; 98JP-0080597.  
XX  
PA (TANA ) TANABE SEIYAKU CO.  
XX  
DR WPI; 1999-626936/54.  
DR N-PSDB; AAZ22701.  
XX  
PT An ABC transporter gene - used to increase the ability of a microbe to  
PT secrete proteins  
XX  
PS Claim 8; Page 21-23; 28pp; Japanese.

QY This sequence represents an ATP-binding cassette (ABC) transporter  
CC component A from a Pseudomonas fluorescens strain 33 and encoded by  
CC the ABC transporter operon II. The operon comprises 3 genes where the  
CC termination codon of the first gene overlaps with the initiation codon  
CC of the second gene. The sequences of the encoded proteins have 63, 56  
CC and 59% homology respectively to the lipBD proteins from Serratia  
CC marcescens. The novel gene and protein can give or increase the  
CC ability of a microbe for secreting a protein.

XX Sequence 580 AA:

Query Match 42.6%; Score 1195; DB 20; Length 580;  
Best Local Similarity 45.5%; Pred. No. 9.7e-94;  
Matches 240; Conservative 106; Mismatches 176; Indels 6; Gaps 3;

QY 28 VFSEFINILAVSPLYMLOYVDRVLTSMNSTLIVLVICVLEFLVYGLLEALRTQVLR 87  
DB 28 lfsavInIlAvpLyMlOyDrVlTsmNsTlIvLvIcVlEflVYgLLeAlRtQvLr 83  
QY 88 GGLKFDGVARDPFKSVYDSTLSRKG---IGGAFRDMDOVREPMTG-GLTAFCDAPWTP 143  
DB 88 lgtqmdmrlngryvdaafeaql--kgytgaagqalsdltlrqfatqgaIfaIfdApWfP 141  
QY 144 VFVIVSMVLPFFGLAIIICIIIFGLAVMMDNAKNPDMQATMASTIAQONDAGSTLRNA 203  
DB 144 vYvIvSmVlpffGLaIIICIIIFglAVmmdnAKNpDmQATmASTIAQondAGsTLrNa 201  
QY 204 EVNRAMGMWGGLQARWRARRDROVMOAASDAGAAVSGIKVFRNIVOTLLIGGAVLA 263  
DB 202 eaiemgmglatlrarwrlaqhgfiaaqlasekaalsawskyrllaqslvlglaIIa 261







Result	6
AAW22153	AAW22153 standard; Protein; 707 AA.
XX	
AC	AAW22153;
XX	
DT	16-FEB-1998 (first entry)
XX	
DE	ApxIB protein.
KW	RTX toxin; apxICA gene; apxIBD gene; apxIAB/C gene; apxIIABCD gene;
XX	repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
XX	therapy; A. pleuropneumoniae infection; swine pleuropneumonia.
OS	Actinobacillus pleuropneumoniae.
XX	
PN	CA2170839-A.
XX	
PD	02-SEP-1996.
XX	
PF	01-MAR-1996; 96CA-2170839.
XX	
PR	01-MAR-1995; 95US-0396244.
XX	
PA	(UYGU-) UNIV GUELPH.
XX	
PI	MacInnes J, Mallard B, Ricciattelli P, Rosendal S;
DR	WPI; 1997-245536/23.
DR	N-PSDB; AAT73218.
XX	
PT	Preparations of microorganisms producing cell-associated RTX toxins
PT	-especially for production of vaccines against swine
XX	pleuro-pneumonia
PS	Disclosure; Pages 93-95; 151pp; English.
XX	
CC	AAW22151-722161 represent A. pleuropneumoniae RTX (repeat in toxins)
CC	toxins. These sequence are encoded by the apxICA, apxIBD, apxIAB/C, and
CC	apxIIABCD genes (see AAT73217-T73220), and can be expressed by
CC	microorganisms used in the preparations of the invention. The
CC	preparations are bacterial preparations comprising one or more isolated
CC	and purified strains of a microorganism that produces one or more RTX
CC	toxins, where the strains have at least one cell-associated RTX toxin.
CC	The preparations are used for production of vaccines for the prophylaxis
CC	and treatment of infectious diseases caused by microorganisms that
CC	produce RTX toxins, where the strains have been attenuated or
CC	inactivated. The vaccines are preferably against Actinobacillus
CC	pleuropneumoniae infection (swine pleuropneumonia). It has been found
CC	that A. pleuropneumonia produces significant quantities of
CC	cell-associated RTX toxins when cultured under certain conditions, and
CC	that the whole-cell protein composition of the cultures corresponds to
CC	the whole-cell protein profiles obtained from cells recovered at
CC	necropsy from the pleural fluid of infected swine. Vaccination with a
CC	bacterin prepared from heat-inactivated cultures having significant
CC	quantities of cell-associated RTX toxins give significant protection of
CC	swine against challenge with homologous strains.
XX	
SQ	Sequence 707 AA;
Query Match	20.3%; Score 569.5; DB 18; Length 707;
Best Local Similarity	28.4%; Pred. No. 4.5e-40;
Matches 161; Conservative 105; Mismatches 253; Indels 47; Gaps	9.
QY	10 TTFDQAVLVARPAVITAMFSEFFIITIALVSPSLMOQYDVRVLTSRNMTLIVLVICVF 69
Db	141 twfipavikykklfietlvisflqifalipcllfqgvmckxlvnrgfstlnhivcalai 200
OY	70 LFVYVGLLEALRTQVLVVGGLKFGDVGARDPRKSVLDSTLS-----RKGIGQAFRDMQ 124
Db	201 vvlfeivnigrltyfahstsrldvelgarlfrhlhlpisvfenrrygdvtvarveidq 260

QY 125 VAEFTG-GLIAPCDAPRTPEVFLVSMHLPFEGILAIITCIIIFGLAVMNDNATNPQ 183  
D 261 Irfnlftgaltstvidlmtfstfifavmuyypaklclvlvgllspflpymqwsifspilrrtd 320  
QY 184 MATMASIAONDAGSTLRMAEVMKAMGMWGLOARMRARREDEVAMQAA---SDAGA 239  
D 321 ekfargadngsfivesvtalntcltalavcpqmntw---dkqlasyvaagfivrtlatl 376  
QY 240 VMSGKVERNLVQTLILLOGAVLAIDKISAGAMIASLIVGRALPDIGAWGMKNYIG 299  
D 377 gqgvgvfgdqkvwmvltlwlghnlvisgdslsgqlclftmlmsqvlapvrlrqlwdfqg 436  
QY 300 ARGAMDRLQTLRE-EKSADDMHLPERPQVLSNEAASILPECAOQPTMRQASFR----- 353  
D 437 vglsvtrlgdvlmspteevgkklalpktdgdl-----fltrlnfrkpa 481  
QY 354 -----IDAGAAVNLVSPSAGKSSLLRGIVGWPCACAVIRLDGYDRIKQMDPEKL 403  
D 482 pylndvnlslsgqgevgivgrrsgsgstlklqrfrybngqvivgdhldaldpml 541  
QY 404 GRHVGVPDIELEFSGVVAONIAFTE-FESQEVTEATLAGVHEMIQSLPMGYDTAIGE 462  
D 542 rrygvavlgdvlhlnsrldnlaldpmpmekivbaaklaaghefiselregylntlvge 601  
QY 463 GGAISGGQRRRLARVFRMPALLVLDPNMSLDQVGVALLMEAMKRLKAARTVIFA 522  
D 602 qgaglsqgqrclatalaralvmpklllfdatsaldyesehlrmhql-ckgrtvlvl 660  
QY 523 THKVNLAQADYIMVINGVISDGE 548  
D 661 ahrlstvnadrllymekqglvegk 686

RESULT 7  
AAV51407  
ID AAV51407 standard; protein; 707 AA.  
XX  
AC AAV51407;  
XX  
DT 05-MAY-2000 (first entry)  
XX  
DE A. pleuropneumoniae clyIB protein.  
XX  
KW RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;  
KW antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;  
KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;  
KW urinary fever; abortion; whooping cough; sleepy foal disease; joint ill,  
KW urinary infection; peritonitis; meningitis; gastroenteritis;  
KW passive immunization; clyIB.  
XX  
OS Actinobacillus pleuropneumoniae.  
XX  
PN US6019984-A.  
XX  
PD 01-FEB-2000.  
XX  
PE 23-DEC-1996; 96US-0772270.  
XX  
PR 01-MAR-1995; 95US-0396244.  
XX  
PA (UYGU-) UNIV GUELPH.  
XX  
PI Mallard B, Rosendal S, MacInnes J, Ricciatti P,  
XX WPI: 2000-146864/13.  
XX DR N-PSDB; AA288585.  
XX  
PT Bacterial preparation comprising microorganisms which produce a member  
PT of the Repeats in Toxins (RTX) family, useful for treating swine  
PT pleuropneumonia, arthritis in swine, shipping fever and abortion in  
PT cattle, and sleepy foal disease -

PS Disclosure: Column 55-60; 96pp; English.  
XX  
CC This invention describes a novel bacterial preparation (I) which  
CC comprises one or more isolated and purified strain(s) of a microorganism,  
CC cultured in tryptone yeast extract (TYE) broth, which produces one or  
CC more RIX toxins (belonging to the family of toxins referred to as Repeats  
CC in Toxins), where the strain(s) have at least one RIX toxin which is  
CC cell-associated. The products of the invention have immunostimulatory,  
CC antitumor, antiinflammatory, antirheumatic and antitubercular activity.  
CC The bacterial preparation may be used as vaccines for the prophylaxis and  
CC treatment of infectious diseases caused by strains of microorganisms  
CC which produce one or more RIX toxins. The infectious diseases are swine  
CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;  
CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping  
CC fever and abortion in cattle; whooping cough, sleepy foal disease or  
CC joint ill (purulent nephritis, arthritis) in foals; septicemia,  
CC polyarthritis and abortion in horses; and urinary infections,  
CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations  
CC may also be used to prepare antibodies which may be used as a means of  
CC passive immunization. This sequence represents the Actinobacillus  
CC pleuropneumoniae clyis protein described in the method of the invention.  
XX  
SQ Sequence 707 AA:

Query Match 20.3%; Score 569.5; DB 21; Length 707;  
Best Local Similarity 28.4%; Pred. No. 4,5e-40;  
Matches 161; Conservative 105; Mismatches 253; Indels 47; Gaps 9;

OY 10 TIFDOAVLVARPAMVMTAFSFFINILALVSPMLQYDRVITSRVNSTLIVLTVICVF 69  
DB 141 twifpavikykrlflevlvsvalqfalftlffgvmdkvlhrgfstlvdvssvali 200  
OY 70 LFLVYGLLEAIRTOVLVRGGLKFDGVARDPFRKSVLDSTLS-----RKIGGGAFRMDQ 124  
DB 201 vllfeivnglrltylfahstsrldvejgarlfrhlalpslyferrvgdtvarvlelq 260  
OY 125 VREMTG-GLIAFDAPMTVPFVIVSWMLHPFGILAIACIIIFGLAVMNDATKMPQ 183  
DB 261 lrrfltgalsvldlafsflfaymwysscllwvlasipafwsafispllrln 320  
OY 184 MATNASTAQNDAOSTLRNAAEVMKAMGMGGLQARWRARDEQVAMQAAA-----SDAGGA 239  
DB 321 ekfargadngsfivesvvalntlkalatvtpqmtlw-----dkqlasyvsagfrytclati 376  
OY 240 VMSGIKVFRNIVOTLLIGGAYLAIIDGKISAGAMIAGSILVGRALAPREGAVGOMKNYIG 299  
DB 377 gqgqvqfqlkvwmvltlwgahlvsgdlisqqlafmllsgqvlaiparltaqldq 436  
OY 300 ARGAMDRLQTMRLR-EKSADDMPLPEPRGVLSAEMASILPPGAOQPTMRQASFR----- 353  
DB 437 vglsvtrlgdvlnsplesyqgklalpelkgdl-----tfrnrlfrykpxda 481  
OY 354 -----IDAGAAVALVPSAAGKSLRLGIVGWPCAGVIRLDGYIKQMDPKL 403  
DB 482 pvlindvnlisqggevlgivgrsgsktlcllqrfyipengqvlldghlalaadpwl 541  
OY 404 GRHVGYLRFODLEFSGVTAONIARFTE-FESQEVTEATLGVHEMIOSLPMGDTAIGE 462  
DB 542 rrrgvavlgdhnvllnrslrdhalaadpymeklvhaaklgaghefiselregyntlyge 601  
OY 463 GGASLSGQRORLALARVFRPALVLDPEPNASLDQGEVALMEAMKRLKAARKTVIFA 522  
DB 602 gqaglsqggrqrlataraalvnpklllfdeatsaldysesehlmrmhqi-ckgrtvlll 660  
OY 523 THKAVLLAQAADYIMYINGVSDSEGE 548  
DB 661 ahrilstvknadrlivmekqgvlvegk 686

RESULT 8  
AAAY27214  
ID AAY27214 standard; Protein: 710 AA.

XX  
AC AAY27214;  
XX  
DT 24-SEP-1999 (first entry)  
XX  
DE Amino acid sequence of N. meningitidis protein ORF39-1.  
XX  
KW Neisseria meningitidis protein; pharmaceutical; vaccine; diagnosis;  
KW bacterial infection; treatment.  
XX  
OS Neisseria meningitidis.  
XX  
PN WC0936544-A2.  
XX  
PD 22-JUL-1999.  
XX  
PF 14-JAN-1999; 99WC-IB00103.  
XX  
PR 09-OCT-1998; 98GB-0022143.  
PR 14-JAN-1998; 98GB-0000760.  
PR 01-SEP-1998; 98GB-0019015.  
XX  
PA (CHIR-) CHIRON SPA.  
XX  
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;  
XX  
DR WP1; 1999-444400/37.  
DR N-PSDB; AAX99136.  
XX  
PT New protein and its nucleotide sequence, useful in vaccines or  
PT diagnostic compositions for treating and/or preventing Neisseria  
PT meningitidis infections  
XX  
PS Claim 4; Page 75; 123pp; English.  
XX  
CC The invention provides proteins (AAY27201-245) from Neisseria  
CC meningitidis (strains A and B) and nucleic acid sequences (AAY99123-167)  
CC encoding the proteins. Compositions comprising the protein, nucleic acid  
CC or antibody specific to the protein are useful as pharmaceuticals, e.g. a  
CC vaccine composition or a diagnostic composition. The composition is also  
CC useful for treating or preventing an infection due to Neisseria  
CC bacteria, especially Neisseria meningitidis.  
XX  
SQ Sequence 710 AA:

Query Match 20.2%; Score 567.5; DB 20; Length 710;  
Best Local Similarity 29.7%; Pred. No. 6,8e-40;  
Matches 167; Conservative 110; Mismatches 244; Indels 41; Gaps 13;

OY 10 TIFDOAVLVARPAMVMTAFSFFINILALVSPMLQYDRVITSRVNSTLIVLTVICVF 69  
DB 144 twifpavikykrlflevlvsvalqfalftlffgvmdkvlhrgfstlvdvssvali 203  
OY 70 LFLVYGLLEAIRTOVLVRGGLKFDGVARDPFRKSVLDSTLS-----RKIGGGAFRMDQ 124  
DB 204 vllfeivnglrltylfahstsrldvejgarlfrhlalpslyferrvgdtvarvlelq 263  
OY 125 VREMTG-GLIAFDAPMTVPFVIVSWMLHP-FFGILAIACIIIFGLAVMNDATKMPQ 174  
DB 264 lrrfltgalsvldlafsflfaymwysscllwvlasipafwsafispllrln 323  
OY 175 DNATKN-PIOMATNASTAQNDAOSTLRNAAEVMKAMGMGGLQARWRARDEQVAMQAAA 233  
DB 324 dkfarnadngsfiveslta-----vgtvkamavepmtqtw-----dnqlaayvas 369  
OY 234 -----SDAGGAVMGSIKVFVRNIVOTLLIGGAYLAIIDGKISAGAMIAGSILVGRALAPREG 289  
DB 370 gfrvcllavvgqgqvqlqklvtvatlwlgarlvleskltyvgqlafmllsgqvaaapvir 429  
OY 290 AVGOMKNY--LIGARGAMDRLQTMRLR-EKSADDMPLPEPRGVLSAEMASILPPGAOQPT 346  
DB 430 laqlwqldqvgvisva--rlgdilnaptenasshlaipdrigtelifehvdtriyadgrlli 487

QY 347 MKQSFRIIDGAAVALVGPSPAAKSSLLKINGVWPCAAAGVIRLDGCDIKQMPDKLRH 406  
 Db 488 IqdnIrlrIgeVglvIrgsgskstlKlvgrLlyvpeggrvIvgnadlaaPawlrTq 547  
 QY 407 VGYLPDQIEFLFSGVAVQONIA-RFTFEPSQEVTEAATLACVHFMQISLPWGCVTAIGEGA 465  
 Db 548 vgvvIlgenvllnsrlchnlaltdtgmplelrlleaKlgahefnlpegyIvvgvgega 607  
 QY 466 SLISGGQRRLALRAVAFRMPALVLDEPNASLDQVGEVALMEAMRLKAARTVIFATRK 525  
 Db 608 gIsgggqgrIataraIrltlnprllIfdeatsaldySeraIImgmgaI-cantvIlaIhr 666  
 QY 526 VNLLAADYIMVINGVYISDFG 547  
 Db 667 IstvtahrlIamdkgrIvIeaq 688  
 RESULT 9  
 ID AAU36153 standard; Protein; 560 AA.  
 AC AAU36153;  
 XX  
 DT 14-FEB-2002 (first entry)  
 DE Klebsiella pneumoniae cellular proliferation protein #141.  
 XX  
 KW Antisense; prokaryotic cellular proliferation protein;  
 XX antibiotic; antibacterial; drug design.  
 OS Klebsiella pneumoniae.  
 XX  
 PN W0200170955-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 21-MAR-2001; 2001MO-US09180.  
 XX  
 PR 21-MAR-2000; 2000US-191078P.  
 PR 23-MAY-2000; 2000US-206848P.  
 PR 26-MAY-2000; 2000US-207727P.  
 PR 23-OCT-2000; 2000US-242578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX  
 DR WPI; 2001-611495/70.  
 DR N-PSDB; AAS54012.  
 XX  
 PT New polynucleotides for the identification and development of  
 PS antibiotics, comprise sequences of antisense nucleic acids -  
 PS  
 XX Example 3; Seq ID No 11746; 511pp; English.  
 CC  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in

a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.  
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at  
[ftp.wipo.int/pub/published\\_pct\\_sequences](ftp:wipo.int/pub/published_pct_sequences).

Sequence 560 AA:

Query Match 19.8%; Score 555.5; DB 22; Length 560;  
Best Local Similarity 30.4%; Pred. No. 5.2e-39;  
Matches 173; Conservative 96; Mismatches 207; Indels 93; Gaps 18;

QY 27 MYVSEFINILAVSPDLYMLOVDYDRVLTSMNSTLVL----TVICVFPLVGLLEALRT 82  
| : | ||||| : : ||||| : : | : | : | : | :  
Db 13 maslaiaihlalaagllfsmgydydrylpqaspyellvylsfgyiafvnf---vlkitrg 68  
QY 83 QVLVWGKLFEDGVANDPIFKSV--LDSTLSRKIGG--QAERDMDOVEEFTGTGLI-AFC 137  
| : | : | : | : | : | : | : | : | : | :  
Db 69 hvtldlgkrgrdiridryfgahrlkhktaevstgsffisqrdleqvremtstmtvia 128  
QY 138 DAPMPVPFVYSWMLHPFGILLAIACIIIF--GLAVNNDNAKTPIQMATMASTAAAND 195  
| : | : | : | : | : | : | : | : | : | :  
Db 129 dlpflilqlvlfvlfspslwslawpaellmlpljlllk-----klaelankn 177  
QY 196 AGSTLRNAV-----MKAMGMGGQLQARWRARDQVAMQAAASDAG-----G 238  
| ||||| : : | : | : | : | : | : | :  
Db 178 hesltinalivesigglpdikmmgeesflgwn-----syaltaesgvqtktvhg 230  
QY 239 AVMSGIKVFRNIIVTLLLEGAYILAIDKISAGAMIAGSILGRALAEI--GAVGOMK 295  
| : | : | : | : | : | : | : | : | : | :  
Db 231 lvswwmsl-qulllyatvaavgaplvlngdiltgmnaasmlsrnvapmatalcylvlatwg 289  
QY 296 NTIGARGAMDRLQTMRREKSADHMPLEP---RGVLSAFAASILPPGAQQPTTKRAS 352  
| : | : | : | : | : | : | : | : | : | :  
Db 290 qvkvakasilnalmapveg-----gidreplhrahvgnv-----efinaef 331  
QY 353 R-----IDAGAVALLVGSAGRSKSSLLRGIVGVMPCAGAVIRLDGYDK 396  
| : | : | : | : | : | : | : | : | : | :  
Db 332 rykgdnpgpeplrikqskakekiaivlgigqskcllqgamngventsegitrldslip 391  
QY 397 QMDPEKLGRHWYLPDIIEFSGIVAQNIAETFEES--OEVI EAATLAGVHEMIOSLPMG 455  
| : | : | : | : | : | : | : | : | : | :  
Db 392 qgdldirtnatlqgarlfbgtlrrentvitgrpnasddelvkvelogalefnkplmg 451  
QY 456 YDTAIEGCASISGCGORQLARAVFRKPALLVIDEPNASIDOVGEVLAEMAKRLKAA 515  
| : | : | : | : | : | : | : | : | : | :  
Db 452 lehlmweglqigsqgqsllaralirdpnlllildptsfdertekafveqlarw-ag 510  
QY 516 KRTVIEATHTKVNLAADQYIMVINGVIS 544  
| : | : | : | : | : | : | : | : | : | :  
Db 511 erlliahckaaavlndrvlviqgqla 539

RESULT 10  
AAW22160  
ID AAW22160 standard: Protein: 711 AA.  
XX  
AC AAW22160;  
XX  
XX DT 16-FEB-1998 (first entry)  
XX DE ApxIIIIC protein.  
XX RTX toxin; apxICA gene; apxIBD gene; apxIIAB'C gene; apxIIABCD gene;  
KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;  
KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia.  
XX Actinobacillus pleuropneumoniae.  
OS  
PN CA2170839-A.

PD 02-SEP-1996.  
 XX 01-MAR-1996; 96CA-2170839.  
 XX 01-MAR-1995; 95US-0396244.  
 XX (UYGU-) UNIV GUELPH.  
 XX MacInnes J, Mallard B, Ricciatti P, Rosendal S;  
 DR WPI, 1997-245536/23.  
 DR N-PSDB; AAT73220.  
 XX Preparations of microorganisms producing cell-associated RTX toxins  
 PT - especially for production of vaccines against swine  
 PT pleuro-pneumonia  
 PS Disclosure; pages 111-113; 151pp; English.  
 XX AAM22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)  
 CC toxins. These sequence are encoded by the apxICA, apxIIB, apxIIAB/C, and  
 CC apxIIABD genes (see AAT73217-R73220), and can be expressed by  
 CC microorganisms used in the preparations of the invention. The  
 CC preparations are bacterial preparations comprising one or more isolated  
 CC and purified strains of a microorganism that produces one or more RTX  
 CC toxins, where the strains have at least one cell-associated RTX toxin.  
 CC The preparations are used for production of vaccines for the prophylaxis  
 CC and treatment of infectious diseases caused by microorganisms that  
 CC produce RTX toxins, where the strains have been attenuated or  
 CC inactivated. The vaccines are preferably against Actinobacillus  
 CC pleuropneumoniae infection (swine pleuropneumonia). It has been found  
 CC that A. pleuropneumonia produces significant quantities of  
 CC cell-associated RTX toxins when cultured under certain conditions, and  
 CC that the whole-cell protein composition of the cultures corresponds to  
 CC the whole-cell protein profiles obtained from cells recovered at  
 CC necropsy from the pleural fluid of infected swine. Vaccination with a  
 CC bacterin prepared from heat-inactivated cultures having significant  
 CC quantities of cell-associated RTX toxins give significant protection of  
 CC swine against challenge with homologous strains.  
 XX Sequence 711 AA;  
 SO

Query Match 19.7%; Score 552.5; DB 18; Length 711;  
 Best Local Similarity 28.9%; Pred. No. 1.3e-38;  
 Matches 165; Conservative 117; Mismatches 238; Indels 51; Gaps 14;

QY 10 TIFDOAVLVARPAVITAMVFSSFINILALVSPLYMLQYDRLVTSRNSTLLVLYVICVF 69  
 DB 145 twfipavikykrlfveclivlsifqlfalfplffgvmdkvlvnhglslnlvtalsv 204  
 QY 70 LFLVVGLEALFTOVLVRGGLKFDGVARDPKFSVDSRLS-----RKIGGGAFFRMDQ 124  
 DB 205 vvlfeivlsqrltylfsststidvlgaklrfhlalpsifentrrvgtvayreldq 264  
 QY 125 VREFMTG-GLIAFCAPMTPEVIVSWMLPPEGLIAITA--CIITGLAVNMNATKMP 181  
 DB 265 lnflfgqaltsevlldlfsfifawmypsplttvllslpcylawsfif-----sp 317  
 QY 182 IGMATMASTAAON-DAGSTLRNA---EVMKAMGMWGIGQARRRRARDQVAMQAAASDA 236  
 DB 318 llrrrldexfarnadnqsfivesvsaidtkalavrpqtnlW---dkqlasysaa-df 372  
 QY 237 GGAAMS-----GIKVRNIVYQTLILGGGAVIAIDGKISGAMFAGSILGRALAPLEGAV 291  
 DB 373 rvtvlatiqgqvglqktvmiinlwlghlvsqslsqqltltmmlsqvlapvr1a 432  
 QY 292 GOMKNYIGARGAMDRLOTMLE-ERKADDMPLPEPRGYLS-----AEASILPBG 341  
 DB 433 qlwqtdqvgvlsitrlgtvlnspenlygklsipeifgdtaefhrrfrykpadpild-- 490  
 QY 342 ADOPTMROASFRIDAGAAVALVGPASAAGSSLLRGIVGWPACAVYIRLDGYDIKMDPE 401

DB 491 -----dnlsvkqgeivivgrsgsketlklldgrfipengqylidghdalaadpn 543  
 QY 402 KLGRRHVGTLPODIELEFGTVAQNIARTEFESEQ-VIEAATLGAHEMISLPMGYDTAI 460  
 DB 544 wlrrqisgvlnqdnvllnrstlrdnaldcpmsmervlyaaaklagahdfiselsegynltiv 603  
 QY 461 GEGGASLSGGORQLALARAFAVRFPALLVDEFPNASLDQGEVALMEAMRLKARRTVI 520  
 DB 604 gelyagisggqrqrlataraalvmprrillldealsaldysesehlmgmqkl-chgrtlvi 662  
 QY 521 FATKVNILAQADYIYVINOGVISDPGERDR 551  
 DB 663 tlahrstcvknadrllvymekyhivegkhng 693

RESULT 11  
 AAY51413  
 ID AAY51413 standard; protein: 711 AA.  
 AC AAY51413;  
 DT 05-MAY-2000 (first entry)  
 XX A. pleuropneumoniae apxIIIB protein.  
 DE  
 XX RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;  
 KW antiinflammatory; antiarthritic; antiabortive; treatment; pneumonia;  
 KW pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;  
 KW shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;  
 KW urinary infection; peritonitis; meningitis; gastroenteritis;  
 KW passive immunization; apxIIB.  
 KW  
 XX Actinobacillus pleuropneumoniae.  
 OS  
 XX US6019984-A.  
 PN  
 XX 01-FEB-2000.  
 PD  
 XX 23-DEC-1996; 96US-0772270.  
 PF  
 XX 01-MAR-1995; 95US-0396244.  
 PR  
 XX (UYGU-) UNIV GUELPH.  
 PA  
 XX Mallard B, Rosendal S, MacInnes J, Ricciatti P;  
 PT WPI, 2000-146864/13.  
 DR N-PSDB; AAZ88587.  
 XX Bacterial preparation comprising microorganisms which produce a member  
 PT of the Repeats in Toxins (RTX) family, useful for treating swine  
 PT pleuropneumonia, arthritis in swine, shipping fever and abortion in  
 PT cattle, and sleepy foal disease -  
 PS  
 PS Disclosure; Column 83-88; 96pp; English.  
 XX This invention describes a novel bacterial preparation (1) which  
 CC comprises one or more isolated and purified strain(s) of a microorganism,  
 CC cultured in tryptone yeast extract (TYE) broth, which produces one or  
 CC more RTX toxins (belonging to the family of toxins referred to as Repeats  
 CC in Toxins), where the strain(s) have at least one RTX toxin which is  
 CC cell-associated. The products of the invention have immunostimulatory,  
 CC antimicrobial, antiinflammatory, antiarthritic and antiabortive activity.  
 CC The bacterial preparation may be used as vaccines for the prophylaxis and  
 CC treatment of infectious diseases caused by strains of microorganisms  
 CC which produce one or more RTX toxins. The infectious diseases are swine  
 CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;  
 CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping  
 CC fever and abortion in cattle; whooping cough, sleepy foal disease or  
 CC joint ill (purulent nephritis, arthritis) in foals; septicemia,  
 CC polyarthritis and abortion in horses; and urinary infections,  
 CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations  
 CC may also be used to prepare antibodies which may be used as a means of

CC passive immunization. This sequence represents the *Actinobacillus*  
CC *pleuropneumoniae* apxiib protein described in the method of the  
CC invention.

XX  
SQ Sequence 711 AA:

Query Match 19.7%; Score 552.5; DB 21; Length 711;  
Best local similarity 28.9%; Pred. No. 1.3e-38;  
Matches 165; Conservative 117; Mismatches 238; Indels 51; Gaps 14;

QY 10 TIRDAVLARAVITAMVFFINILALVSPYLMQVDRVLSRNVSLVLYVICVF 69  
DB 145 twfipavikyrkifveitlvsiflqlfalltpiffyvmakvlvnhgfcslvltvalsv 204  
QY 70 LFLVYGLLEALRQVLYVRLGKFDGVARDPFEKSVLDSTLS-----RKGGGAFRMDQ 124  
DB 205 vvfefvlsqirfyfshstsrldvelgaktfhlhllpisyfentrvgtvdvareldq 264  
QY 125 VREFMTG-GLIAFCDAWPTVEVIVSMMLHPEFGILALIA--CIIFGLVAMNDNAKPN 181  
DB 265 lrrflfggltsvidlffiffavmypsaktivllslpoylavsifi-----sp 317  
QY 182 IQMATMASTIAON-DAGSTIRNA-----EVKAKGMNGGLQARRRARRDEQVAMQAAASDA 236  
DB 318 lrrrldekfarnadnqsfivesaidsitkavtprqmtulw----dkqlasyvsa-df 372  
QY 237 GGVVMS-----GIKVRNRYQTLILGGAYLADSKISAGAMTAGSTIVGRALAPLEGAV 291  
DB 373 rvtvltlqgqgqyqlkqtwminlwlghnlvsqdsisqqltftmllsqvlapvrila 432  
QY 292 GOMKNYIGARGMDRLQTMLE-EKSADDMPLPEPRGVLS-----AEAASILPPG 341  
DB 433 qlwqdfqgqyisltirgvdinsptenygqlslpelgdfafhifryqpdpilld-- 490  
QY 342 AQQPTMRQASFRIDAGAAVALVGPSSAGSSSLRGIVGWPCAAGYIRLDGYDIKQMDPE 401  
DB 491 -----dvnlsykvgevgivgrsgsgsktlklqlrfyfpengvylldghdlaadpn 543  
QY 402 KIGRHGYVLPDDELFGSTVMAQNIARFTEFESQE-VTEAATLIGVHEMIOSLMGVDTAI 460  
DB 544 wlrtrqgvvldqnvllnrsirtdialtdpsmsmervlylaaklsagadfsellegyntliv 603  
QY 461 GEGGASLSSGSGORLALARAFVRMPALLVDEPNASLDQGEVALMEAMRLKAARKTVI 520  
DB 604 gelgaglsqgqgrlatarelavmnprrlllfdestalsdyesehlingmqki-chgrtvi 662  
QY 521 FATHKYNLLAODYIWINOGVISDFGERDR 551  
DB 663 llahtistvknadrlivmekghivqegkhng 693

RESULT 12  
AA033597  
ID AAU33597 standard; Protein; 723 AA.  
XX  
AC AAU33597;  
XX  
DT 14-FEB-2002 (first entry)  
DE Pseudomonas aeruginosa cellular proliferation protein #41.  
XX  
KW Antisense; prokaryotic cellular proliferation protein;  
XX antibiotic; antibacterial; drug design.  
OS Pseudomonas aeruginosa.  
XX  
PN WO200170955-A2.  
XX  
XX 27-SEP-2001.  
PD 21-MAR-2001; 2001WO-US09180.  
XX  
XX

PR 21-MAR-2000; 2000US-191078P.  
PR 23-MAY-2000; 2000US-206848P.  
PR 26-MAY-2000; 2000US-207727P.  
PR 23-OCT-2000; 2000US-242578P.  
PR 27-NOV-2000; 2000US-253625P.  
PR 22-DEC-2000; 2000US-257931P.  
PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELIRRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
PI WPI; 2001-611495/70.

DR N-PSDB; AAS51456.

PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 5093; 51pp; English.

CC The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the  
CC genes, their use in the discovery of novel antibiotics, the essential  
CC genes themselves and the encoded proteins. The prokaryotes used are  
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
CC invention is also useful for the identification of potential new targets  
CC for antibiotic development. The antisense nucleic acids can also be used  
CC to identify proteins used in proliferation, to express these proteins,  
CC and to obtain antibodies capable of binding to the expressed proteins.  
CC The proteins can be used to screen compounds in rational drug discovery  
CC programmes. The antisense nucleic acid sequence is also useful to screen  
CC for homologous nucleic acids which are required for cell proliferation in  
CC a wide variety of organisms. The present sequence represents an  
CC essential prokaryotic cellular proliferation protein.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 723 AA:

Query Match 19.6%; Score 549.5; DB 22; Length 723;  
Best local similarity 29.3%; Pred. No. 2.4e-38;  
Matches 160; Conservative 106; Mismatches 241; Indels 39; Gaps 14;

QY 22 AVITAMVFFINILALVSPYLMQVDRVLSRNVSLVLYVICVFLVYGLLEALR 81  
DB 170 aevtae--alvanvlasasalifamgydyrvvnaafdlwlasgvalavldgvlmr 227  
QY 82 TQVLYVRLGKFDGVARDPFEKSVLDSTLSRKGIIGQAF-----RDMQVREFMTGGLIA-F 136  
DB 228 ghlinvlgkrldlqstllsrvtvaakpsmgfscvqefesvreflssaa1l 287  
QY 137 CDAPMPVPEVIVSMMLHPEFGILAIACITIF--GLAVMNDNAKTNFIQMATMASTIAON 194  
DB 288 sdipfvafillilavlgghvwwpvlvacvlllpgl-----lqrlighsrpn 336  
QY 195 DAGSTLRNAEVRKAMGMLQARRRARRDEQVAMQAAASDAGAVW-----SGIKVFR 248  
DB 337 lrgamkngyllaefelhlevkat-rtegcclhngwellgelgatkthltlastlsysa 395  
QY 249 NIYQTLILG-----GAYLAIDKISAGAMTAGSTIVGRALAPIE--GAVGOMKNYIGAR 301  
DB 396 slvqqlcygvvuvfygryrisesgamtvglvacslsaratapsqagllgrvqhtkvam 455  
QY 302 GAMDRLQTMREKSSADDHMPRLPEPRGVLSAEASILPPGAQQPTMRQASFRIDAGAVA 361  
DB 456 egljdlmsaeqepqgkrrfthkerlqghyrlegvr--lahngdspvvavqalntraegera 514  
QY 362 LVGPSAAGKSSSLRGIVGWPCAAGYIRLDGYDIKQMDPEKGRHVGYLPDIELEFSGTV 421

Db 515 llygnagagstllrlrlslglladaqgllldvslctqdpadgrgrylgylpqdvalthgs1 574  
Qy 422 AONI-ARFTEFESQEVIEATTAAGVHEMTQSLPMGYDTAIGEGGASLGGORRLALARA 480  
Db 575 rdhlnlenaalgdeellelctldvgylgaftvgrhplgidmptl-gqnalslsgqraqavglarv 633  
Qy 481 VFEMPALVLDPEFNASLDQVGEVALMEAMKRLKAARTRVFATRKVNLLAQAADYIWINO 540  
Db 634 llydppillllldeptaafdgsgkqvldylyqgw-lykrtllvtlthkksmlalveravlyrq 692  
Qy 541 G-VISD 545  
Db 693 gvyimd 698  
RESULT 13  
AAU38443  
ID AAU38443 standard; Protein; 718 AA.  
XX AC AAU38443;  
XX DT 14-FEB-2002 (first entry)  
XX DE Salmonella typhi cellular proliferation protein #334.  
XX KM Antisense: prokaryotic cellular proliferation protein;  
XX KM antibiotic; antibacterial; drug design.  
XX OS Salmonella typhi.  
XX PN WO200170955-A2.  
XX PD 27-SEP-2001.  
XX PF 21-MAR-2001; 2001MO-US09180.  
XX PR 21-MAR-2000; 2000US-191078P.  
XX PR 23-MAY-2000; 2000US-206848P.  
XX PR 26-MAY-2000; 2000US-207727P.  
XX PR 23-OCT-2000; 2000US-242578P.  
XX PR 27-NOV-2000; 2000US-253625P.  
XX PR 22-DEC-2000; 2000US-257931P.  
XX PR 16-FEB-2001; 2001US-269308P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
XX PI Yamamoto RT, Xu HH;  
XX DR WPI: 2001-611495/70.  
XX DR N-PSDB; AASS6302.  
XX PT New polynucleotides for the identification and development of  
XX PT antibiotics, comprise sequences of antisense nucleic acids -  
XX PS Example 3; Seq ID No 14036; 511pp; English.  
XX CC The invention relates to antisense inhibitors of genes essential to  
XX CC prokaryotic cellular proliferation, their use in identifying the  
XX CC genes, their use in the discovery of novel antibiotics, the essential  
XX CC genes themselves and the encoded proteins. The prokaryotes used are  
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
XX CC pneumoniae, pseudomonas aeruginosa and Enterococcus faecalis. The  
XX CC invention is also useful for the identification of potential new targets  
XX CC for antibiotic development. The antisense nucleic acids can also be used  
XX CC to identify proteins used in proliferation, to express these proteins,  
XX CC and to obtain antibodies capable of binding to the expressed proteins.  
XX CC The proteins can be used to screen compounds in rational drug discovery  
XX CC programmes. The antisense nucleic acid sequence is also useful to screen  
XX CC for homologous nucleic acids which are required for cell proliferation in  
XX CC a wide variety of organisms. The present sequence represents an  
XX CC essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 718 AA.  
Query Match 18.6%; Score 521.5; DB 22; Length 718;  
Best Local Similarity 28.6%; Pred. No. 5,9e-36;  
Matches 164; Conservative 118; Mismatches 221; Indels 71; Gaps 19;  
Qy 8 KPTIEDAVLVARPAVITAVESFETINILALVSPYMLQYIVDRVTSRNVSTIYIVTVC 67  
Db 153 kpdwmxrelvlgdlrpylpwvaafllnvslagvlfsmgydvdrvpagsypltylvls-ftg 211  
Qy 68 VFLEFLVYG-LLEALRFOYLVRGGLKFDGVARDIPFSVL--DSTLSRK-GIGQAFRDM 122  
Db 212 vlvaavlfglfllrearthlndvlygkradmrsdrvlghallrlmsaiprstgstfsglrel 271  
Qy 123 DQVREFMTGGLIA-FCDAPWTPVFVIVSWMLHPFGLIATACIIIF-GLAVMNDNRK 179  
Db 272 eqdremitselctivdlpftflfmvlaliaplrawlapvaallmlypsvalqk----- 326  
Qy 180 NPTOMATMASIAAONDAGSTLRNA-----EVMKA-----MGW-----GG 214  
Db 327 ---klavlangahe---alrnaivesvgjlediklmgaerflqgwmsyrlrtgseg 380  
Qy 215 LQARWRARRDEQVAMQAAADAGAVMSGIFKVFNRNIYQTLILGGAVLATIDGKISAGAMI 274  
Db 381 lrtklkltr--llsw-----gmsv-gslvaavimfgapmvlegsmgtgavv 424  
Qy 275 AGSILVGRALAPIE---GAVGOMKNYIGARGAMDRLQTMLEKESADDMPLPEPRGVLS 331  
Db 425 aaemlgsrmlapmanlcvlarwqvkaakmgldnlnqldetqhdsl1hrd1lbgny1 484  
Qy 332 AEAASIT-LPPGAQOPTMRQASFRIDAGAAVALVPSAAGSSSLRGIVGWPACAGVRL 350  
Db 485 fenaqfryhnddqrllrllrvleimpgerlallgrnagqstllqmaggllem1qgarl 544  
Qy 391 DGYDIKQMDPEKIGRHGYVLPDIELFSGTVAQNIATFEESOE-VIEATTAAGVHEMI 449  
Db 545 dnslshldmadlrmnlgflsgnarllfvgclrenlllgaphandeqlfdalevsqgavfv 604  
Qy 450 QSLPMGYDTAIGEGGASLGGORRLALARAVERFPALVLDPEFNASLDQVGEVALMEAM 509  
Db 605 rrlakgldhplmegngnlgsgqrgslllarmllrspiavilldepsasidhte---refl 661  
Qy 510 KRLK--AAKRTVIFATRKVNLLAQAADYIWINOG 541  
Db 662 qrlhqwlygnrtlvavthrvplllelvervvvlkeg 695  
RESULT 14  
AAU39313  
ID AAU39313 standard; Protein; 571 AA.  
XX AC AAU39313;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polypeptide SEQ ID NO 2458.  
XX KM Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX KM peripheral nervous system; neuropathy; central nervous system; CNS;  
XX KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
XX KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;  
XX KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
XX KM Leukaemia.  
XX OS Homo sapiens.  
XX PN WO200153312-A1.



XX	26-JUL-2001.
PD	
XX	
XX	26-DEC-2000; 2000MO-US34263.
PF	
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PR	09-JUL-2000; 2000US-0596042.
PR	19-JUL-2000; 2000US-0620312.
PR	03-AUG-2000; 2000US-0653450.
PR	14-SEP-2000; 2000US-0662191.
PR	19-OCT-2000; 2000US-0693036.
PR	29-NOV-2000; 2000US-0727344.
PA	
XX	(HYSEQ- ) HYSEQ INC.
PI	Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI	Zhao Q, Zhou P, Goodrich R, Drmanac RT;
XX	
DR	WPI: 2001-442253/47.
XX	N-PSDB: AAI58469.
PT	Novel nucleic acids and polypeptides, useful for treating disorders.
PT	Such as central nervous system injuries -
PS	Example 4: SEQ ID NO 2458; 10078pp; English.
XX	
CC	The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	centralised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
CC	
XX	Sequence 571 AA:
SQ	
Query Match	16.3%; Score 457; DB 22; Length 571;
Best Local Similarity	27.3%; Pred. NO. 1.5e-30;
Matches 153; Conservative 102; Mismatches 234; Indels 72; Gaps 20;	
33	INITIALVSLVLMQYVDRVTSRNVSTLVLTVICVLEFLVY-----GLLEALRQ 83
7	INVLV--PIFYRNVN--LLENAPWNSLWLVESYFLIKLIGGGTGTGVSNIITF 61
84	VLVRLGLKFDGVARDPREFSVLDSLTRKGIQ---GQAFRMDQVREBFTGGL--IAECD 138
62	LVIR-vqgftrrveallfshhelslwhysrrtgevlridrtstvtgllsyivlv 120
139	ADWTEVFI---VSMWLHPFFGIATIAICIIIFGLA-VMNDNATKNPIOMATMASIAQ 193
121	IPCLIIIGLIIIGLYSFMFFNAWGLIVFLQMSYILCLTVVTEWRKLFRRAMTGEN-ATR 179
194	NDAGSTLRNAEYPMKAMGMWGIQIAWRAR--RDEOVANQAAASDAGAVMSIKYFRNIV 251
180	ARAVSLINLFETVKYNAESYEVERYREALIKYGGLEWKSAS-----LVllngt 229
252	CTLIIGGG-----AYLAIDKISAGAMINGSILVGCALAPICGAVGQW----- 294
230	qnlvtglvgllagsllcayfvteqklqvvgdyvlfqcyllqlympl-----nwfgelyymiq 284
295	KNYIGARGMDRLQTMLEEKSAAD--HMPLEBPRGVLSAEASILPFGAQOPTMRQAS 351

Db 285 trlfilmemfid---llkeetevklpgagprrlrfkgrliefenvhfsyadgre-rlqdvS 339  
 Oy 352 FRIDAGAAVALPSPSAGKSSLRGIVGWPCCAAGVIRLDQVDIKQOMPEKIGRNVGLP 411  
 Db 340 fcvmpgqclalvpsagagskstllrllfftydssgclldqddsqvtqaslrshgvvp 399  
 Oy 412 ODIELEFSGTVAONI--ARTEFESEOVIEAMTLACVHEMIOSLPKGYDTAIEGGASLSG 469  
 Db 400 gctvlfnldtladhllygrvta-gndevaaagaaglnhalmafegyrctgyegrlkssg 458  
 Oy 470 GQRRLALARAVFRMPALIVDEPNASLDQGEVALMEAMRLKAARFTVIFATHKVNL 529  
 Db 459 gekqvvalrtllkagpjlilldeatsaldtsneraigslakv-canrtllvvalrlstlv 517  
 Oy 530 AOADYIMVINGVISDFGERD 550  
 Db 518 vnadqlvlakdgclvergrhe 538

RESULT 15  
 AAU36908  
 ID AAU36908 standard; Protein: 578 AA.  
 XX  
 AC AAU36908;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 DE Staphylococcus aureus cellular proliferation protein #1078.  
 XX  
 KW Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 WO200170955-A2.  
 PN  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX  
 PI 21-MAR-2001; 2001WO-US09180.  
 PF  
 XX  
 PR 21-MAR-2000; 2000US-191078P.  
 PR 23-MAY-2000; 2000US-206848P.  
 PR 26-MAY-2000; 2000US-207727P.  
 PR 23-OCT-2000; 2000US-242578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX  
 DR WPI: 2001-611495/70.  
 DR N-PSDB: AAS54767.  
 XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 PS Example 3; Seq ID No 12501; 511pp; English.  
 XX  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen  
CC for homologous nucleic acids which are required for cell proliferation in  
CC a wide variety of organisms. The present sequence represents an  
CC essential prokaryotic cellular proliferation protein.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 578 AA;

## Query Match

16.2%; Score 454; DB 22; Length 578;

Best Local Similarity 24.8%; Pred. No. 2,7e-30;

Matches 143; Conservative 102; Mismatches 244; Indels 88; Gaps 14;

```
QY 22 AVITAMWSPFINILALVSPYLMQVDRVL-----TSRWSTLIVLTVCVLEFLV-- 73
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 18 attlvgalktqiml---pilkyaigyvinhahltdekvhhltaigiallftvlvr 74
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 74 -----YGLEALRTOVLVRCGLKEDGVARDPIFKSVLDS 107
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 75 ppleftrqyiaqwtssnklllydirkklynhhgaalsarifa-----nngvqg 119
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 108 TLSRKIGGQAFRDMDOVREPMTGSLAFCDAPWTP-VFVIVSMMLHPFGILAIACII 166
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 120 vlsr-----vindveqtkdflleglmf---vldcftllialsimffldvkltaalf 169
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 167 IFGLAVNMNDNATKNPIOMAT-----MASIAONDAGSTLRNAEVKAMGM-----WGGL 215
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 170 lfffyiltvyffgrlrlkltrersqalaevygfihervqgisvkvstaledneaknfdk 229
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 216 QARWRARDEQVAAQAASDAGAVMSGKVFERNIVOTLILGGAYLAIDKISAGAMIA 275
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 230 ntnfltralkhtrwnays-----faaintvtdigpiivgayaiaisgslvtglaa 282
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 276 GSILVGRALPIEGAVGQWKNYIGARGAMDRLQTMREE--KSADHMPLPPEPRGVLSA 332
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 283 fvygllellfgrlrrlvasftltqsfasmdrvfglidedydkngvgagpieikgrtdl 342
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 333 EAASILPFGAQQPTMRQASFRIDAGAAVALVSPNAGKSSLLRGIVGVWPCAGVIRLDG 392
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 343 dhvsfygndneapilklnlsiekgetvatfvmesggkstllnlprrfydvtsqqlldg 402
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 393 YDIKQWPEKLGRIHVGVLPODIELEFSGTVAQNI--ARTEPESQEVITPATLAGVHEMIQ 450
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 403 hmlkdfltslrnqiglvqgnllfstlvkenlllgrptacd-eevveaakmanahdfim 461
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 451 SLPMGYDTAIGEGASISGGQRRLAARAVFRMPALLVLDENASLDQVGEVALMEAMK 510
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 462 nlpgydtvevgergvklsggqkgrlsariflfnppllldcatsaldseslgeald 521
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 511 RLKAAKRTVITATHKVNLLAQADYIMVINOQVISDFG 547
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 522 vl-skdrllivahrllstlhackivlenghivetyg 557
    | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
```

Search completed: July 19, 2002, 10:43:17

Job time: 137 sec



Db 23 TVGIFTAFINLMLVPSIYMLQVYDRVLPBSRNETITLMLTLMGFMGMSLLEYRSWV 82  
QY 85 LVKGLKFDGVARDPIFKSVLDSTLSR-KGIGGAQAFDMQVREPMTG-GLIAFCDAQWPT 142  
Db 83 VIRIGSOLDMLRLTRVYTAAYEANLNKSSDAQOMLSDLTNLQFLTGLSALFAFFDAPWF 142  
QY 143 PVEVIYSWMLHPFEGILATIACIIIFGLAVMNDATKNPIQMTMTASIAQNDAGSTLRN 202  
Db 143 PLYLVIFLFPNMLGLFALVGAALLLALAVINEVSKRPLGEASKLSIMSGNLASTMLRN 202  
QY 203 AEVWKAMGMGGLQARRARDEQVAMQAAASDAGAVMSGIFYFRNIVOTLLIGGAYL 262  
Db 203 AEVIEALGMLPNLKRFRFGLHQRFLNSQRLASERASRYTITKFRVMSLQSLVGLGGWL 262  
QY 263 AIDGKISAGAMIGSILVGRALPIEGAVGOMKNYIGARGAMDRLOTMLREKSSADHMP 322  
Db 263 AIDGHTIPGMAMIGSILMGTLPADIEQVINWMSYSAKLSYGLVLLLETHPQRTGMS 322  
QY 323 LPREGVLSAEASILPFGAO-QPTMRQASFRIDAGAAVALVGSAGKSSLLRGIVGW 381  
Db 323 LPREGVLSVEGYATPFGSKGDAVLHNVSFAIQPGDVLGIGPSAGKSTLARLVLGIW 382  
QY 382 PCAAGVIRLDGYDKOMDPKLGHNHGYLPQDILFSGTVAQNTARTTEPSOEVIAT 441  
Db 383 PVSEGIYRLDNADLYQNNKDELGYIGYLPQDILFAGTIAENIARNDIDSEKVIDAAK 442  
QY 442 LAGVHEMIOSLPMGYDTAIGEGASLSGQORLALRAVFRMPALLVLDEPNASLDQVG 501  
Db 443 LAGVHEILRPNGYDVSIVNGAGLSGGQKORIGLARALYGDPALVYLDPEPNSLDDAG 502  
QY 502 EVALMEAMKRLKAARKRYVIFATHKVNLLAQADYIMVINOYISDFG 547  
Db 503 EKALNOAIMFLKQRNKTVLTHRTNLSMTSKLLLVNGVNAFG 548

RESULT 2  
US-08-620-605D-2  
; Sequence 2, Application US/08620605D  
; Patent No. 5846811  
; GENERAL INFORMATION:  
; APPLICANT: SHIBATANI, TAKEJI  
; APPLICANT: AKATSUKA, HIROYUKI  
; APPLICANT: KAWAI, ERI  
; TITLE OF INVENTION: GENE PARTICIPATING IN THE MECHANISM OF  
; TITLE OF INVENTION: SECRETION OF ESTERASE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH  
; STREET: PO BOX 747  
; CITY: FALLS CHURCH  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/620.605D  
; FILING DATE: 22-MAR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SVENSSON, LEONARD R  
; REGISTRATION NUMBER: 30,330  
; REFERENCE/DOCKET NUMBER: 0020-3955  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)-205-8000  
; TELEFAX: (703)-205-8050  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 588 amino acids  
; TYPE: amino acid

; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-620-605D-2

Query Match 41.8%; Score 1173.5; DB 2; Length 588;  
Best Local Similarity 44.7%; Pred. No. 5.9e-104;  
Matches 235; Conservative 111; Mismatches 177; Indels 3; Gaps 3;

QY 25 TAVSFSPFINLALVSPLYMLQVYDRVLTSLIVLIVICVFLVLYGLLEAARTQV 84  
Db 23 TVGIFTAFINLMLVPSIYMLQVYDRVLPBSRNETITLMLTLMGFMGMSLLEYRSWV 82  
QY 85 LVKGLKFDGVARDPIFKSVLDSTLSR-KGIGGAQAFDMQVREPMTG-GLIAFCDAQWPT 142  
Db 83 VIRIGSOLDMLRLTRVYTAAYEANLNKSSDAQOMLSDLTNLQFLTGLSALFAFFDAPWF 142  
QY 143 PVEVIYSWMLHPFEGILATIACIIIFGLAVMNDATKNPIQMTMTASIAQNDAGSTLRN 202  
Db 143 PLYLVIFLFPNMLGLFALVGAALLLALAVINEVSKRPLGEASKLSIMSGNLASTMLRN 202  
QY 203 AEVWKAMGMGGLQARRARDEQVAMQAAASDAGAVMSGIFYFRNIVOTLLIGGAYL 262  
Db 203 AEVIEALGMLPNLKRFRFGLHQRFLNSQRLASERASRYTITKFRVMSLQSLVGLGGWL 262  
QY 263 AIDGKISAGAMIGSILVGRALPIEGAVGOMKNYIGARGAMDRLOTMLREKSSADHMP 322  
Db 263 AIDGHTIPGMAMIGSILMGTLPADIEQVINWMSYSAKLSYGLVLLLETHPQRTGMS 322  
QY 323 LPREGVLSAEASILPFGAO-QPTMRQASFRIDAGAAVALVGSAGKSSLLRGIVGW 381  
Db 323 LPREGVLSVEGYATPFGSKGDAVLHNVSFAIQPGDVLGIGPSAGKSTLARLVLGIW 382  
QY 382 PCAAGVIRLDGYDKOMDPKLGHNHGYLPQDILFSGTVAQNTARTTEPSOEVIAT 441  
Db 383 PVSEGIYRLDNADLYQNNKDELGYIGYLPQDILFAGTIAENIARNDIDSEKVIDAAK 442  
QY 442 LAGVHEMIOSLPMGYDTAIGEGASLSGQORLALRAVFRMPALLVLDEPNASLDQVG 501  
Db 443 LAGVHEILRPNGYDVSIVNGAGLSGGQKORIGLARALYGDPALVYLDPEPNSLDDAG 502  
QY 502 EVALMEAMKRLKAARKRYVIFATHKVNLLAQADYIMVINOYISDFG 547  
Db 503 EKALNOAIMFLKQRNKTVLTHRTNLSMTSKLLLVNGVNAFG 548

RESULT 3  
US-08-772-270A-4  
; Sequence 4, Application US/08772270A  
; Patent No. 6019984  
; GENERAL INFORMATION:  
; APPLICANT: MacInnes, Janet  
; APPLICANT: Ricciardi, Paul  
; APPLICANT: Mallard, Bonnie  
; APPLICANT: Rosendal, Soren  
; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR  
; TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bereskin & Parr  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/772.270A

1 FILING DATE: December 23, 1996  
 2 CLASSIFICATION: 424  
 3 ATTORNEY/AGENT INFORMATION:  
 4 NAME: Gravelle, Michelle  
 5 REGISTRATION NUMBER: 40,261  
 6 REFERENCE/DOCKET NUMBER: 6580-81  
 7 TELECOMMUNICATION INFORMATION:  
 8 TELEPHONE: (416) 364-7311  
 9 TELEFAX: (416) 361-1398  
 10 INFORMATION FOR SEQ ID NO: 4:  
 11 SEQUENCE CHARACTERISTICS:  
 12 LENGTH: 707 amino acids  
 13 TYPE: amino acid  
 14 STRANDEDNESS: single  
 15 TOPOLOGY: linear  
 16 MOLECULE TYPE: protein  
 17 ORIGINAL SOURCE:  
 18 ORGANISM: *Actinobacillus pleuropneumoniae*  
 19 IS-08-772-270A-4

Query Match	20.3%	Score 569.5	DB 3	Length 707
Best Local Similarity	28.4%	Pred. No. 5.9e-46		
Matches 161; Conservative	105	Mismatches 253	Indels 47	Gaps 9

[illegible]

RESULT 4  
US-08-772-270A-12  
Sequence 12, Application US/08772270A  
Patent No. 6015984  
GENERAL INFORMATION:  
APPLICANT: MacInnes, Janet  
APPLICANT: Riccietalli, Paul

1 APPLICANT: Mallard, Bonnie  
 2 APPLICANT: Rosendal, Soren  
 3 TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR  
 4 TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES  
 5 NUMBER OF SEQUENCES: 14  
 6 CORRESPONDENCE ADDRESS:  
 7 ADDRESSEE: Bereskin & Parr  
 8 STREET: 40 King Street West  
 9 CITY: Toronto  
 10 STATE: Ontario  
 11 COUNTRY: Canada  
 12 ZIP: M5H 3Y2  
 13  
 14 COMPUTER READABLE FORM:  
 15 MEDIUM TYPE: Floppy disk  
 16 COMPUTER: IBM PC compatible  
 17 OPERATING SYSTEM: PC-DOS/MS-DOS  
 18 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 19 CURRENT APPLICATION DATA:  
 20 APPLICATION NUMBER: US/08/772,270A  
 21 FILING DATE: December 23, 1996  
 22  
 23 CLASSIFICATION: 424  
 24  
 25 ATTORNEY/AGENT INFORMATION:  
 26 NAME: Gravalle, Micheline  
 27 REGISTRATION NUMBER: 40,261  
 28 REFERENCE/DOCKET NUMBER: 6580-81  
 29 TELECOMMUNICATION INFORMATION:  
 30 TELEPHONE: (416) 364-7311  
 31 TELEFAX: (416) 361-1398  
 32 INFORMATION FOR SEQ ID NO: 12:  
 33 SEQUENCE CHARACTERISTICS:  
 34 LENGTH: 711 amino acids  
 35 TYPE: amino acid  
 36 STRANDEDNESS: single  
 37 TOPOLOGY: linear  
 38 MOLECULE TYPE: protein  
 39 ORIGINAL SOURCE:  
 40 ORGANISM: Actinobacillus pleuropneumoniae  
 41  
 42 US-08-772-270A-12

Query Match	19.7%	Score	552.5	DB	3	Length	711
Best Local Similarity	28.9%	Pred. NO.	2.5e-44				
Matches 165	Conservative	117	Mismatches	238		Indels	51
						Gaps	14

```

0Y 10 TLEDDAVLVARPAVITANVSEFEFINIALVALSPLYMLOYDVLTSRNVSTLVLTVICVF 69
    | | | | | : : : : : | : : : : | : : : : |
Db 145 TWEIDAVIKRKRIPEVETIIVISIFQLFALTRIPLEQYVMDKVLHVRGSTLNVITVALSV 204
    | | | | | : : : : : | : : : : | : : : : |
0Y 70 LFLVYGLLELFTOVLVNGLKEFDGVARDPIKSVLDSTLS----RKGIGGAQFMDQ 124
    | : | : | : | : | : | : | : | : | : | : |
Db 205 VVIFEIVLSTGRTRYTFESHSTRIDEVLGAKLFRRLLALPISFENRRKVDTVARRELDQ 266
    | : | : | : | : | : | : | : | : | : | : |
0Y 125 VREPMTG-GLIAFCDAPRTPPVVIVSMNLHFFEGILAIIA--CIITIGLAVMNDATKNP 181
    | : | : | : | : | : | : | : | : | : | : |
Db 265 IRNELTGALTSVLDLFTSFEEFVMMWYSKPLTIVILLSPCYIAWSIFI-----SP 317
    | : | : | : | : | : | : | : | : | : | : |
0Y 162 IQAMTMAITANQ-DAGTGLNNA----EVMAMGMWGGLARMRARDQVAMQAAASDA 236
    | : | : | : | : | : | : | : | : | : | : |
Db 318 ILRRRLDERFARMANDNOSFLVESASIDITIKALVTPQMTNIM---DKOLASVSA-DF 372
    | | | | | : | : | : | : | : | : | : | : |
0Y 227 GGAVWS-----GIKFRNIYQTLTLLGGAYLADIGKISAGAMINGSLVLRALAPIEGAV 291
    | : | : | : | : | : | : | : | : | : | : |
Db 373 RVYVLTATIGGGVODILOQTVMIINLMIGAHVLVIGDLSIGOLITFENMLSGQVAPVRLA 432
    | : | : | : | : | : | : | : | : | : | : |
0Y 292 GQWKNYICARGAMDLEQTMLE- EKXSADHMRLEPRPGVLS-----AEASLTPPG 341
    | : | : | : | : | : | : | : | : | : | : |
Db 433 QLMODEQOVQGISITRLGVLNLSPEENTOGKSLDEPIFGDIAFKIRIRRYKPDATIID-- 490
    | : | : | : | : | : | : | : | : | : | : |
0Y 342 AAOPTMRQASRIDAGAVANVALVSPAGKSSLLGCIAGVWPCAGAVIRLGDYDKQMDP 401
    | : | : | : | : | : | : | : | : | : | : |
Db 491 -----DVNLSTVKGQEVITGVIGSSGSKSTLTRLIQRFTYIPRENGVVLIDGHDLADPN 543
    | : | : | : | : | : | : | : | : | : | : |
0Y 402 KLGRRHVGLPODIELFSGSTVANOIRARTEPESOE-VIEAATLAGVHMIOSLPWGYDTAI 460

```

DB 544 WLRQGVVLDQDVLNRSIRDNIALTPDSMSKERYIAAKLAGADPFLSELRGNTIY 603  
QY 461 GEGGASISGGORLALARAVERFPALLVDEPNASLDQGEVALMEAMRLKAARTVI 520  
DB 604 GELGAGLSGGORRIARAALVNNPRILIFDEATSLADYSEHIIQNMOKI-CHGRTVI 662  
QY 521 FATHKVNLLAQAADYIMVINGVISDFGERDR 551  
DB 663 IIAHRLSTVKNADRIIYMERGHIVEOGKHNO 693

RESULT 5  
US-08-895-522-3  
; Sequence 3, Application US/08895522  
; Patent No. 5858719  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Shah, Puryi  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE  
; TITLE OF INVENTION: TRANSPORT PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/895,522  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0336 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 694 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Genbank  
; CLONE: 1167982  
; US-08-895-522-3

Query Match 14.3%; Score 401.5; DB 2; Length 694;  
Best Local Similarity 25.1%; Pred. No. 7.2e-30;  
Matches 143; Conservative 113; Mismatches 233; Indels 81; Gaps 19;

QY 23 VITANVESFFIITIALVSEPLWLDQYDRVLSRNSTLIVLVTVFLVYGLLEALRT 82  
DB 97 IIVPEMFYAVDSLQMSG-NMLNLSDAPNVATMATAVLIG-----YG----- 139  
QY 83 QVIVAGGLKPGVDARDPLEKFSVLDSTLSKKGIGGQAF-----RMDQVREFMVGGLIATC 137  
DB 140 --VSHAGAAFFNEVKNNAVGVKVAONSIRR--IAKNVFLHLHLNDLGFHLSTGTGALSRAI 195

QY 138 DAPWTEPVIVS---WMHPREFIILAIACILIEGLAVMNDNAKTNPIOMATWASIA--- 191  
DB 196 DRGTGRISEFLSALVNLNLPYEFEMMLVSSVLTKCGAOPALVTLGTAITAFYAVTR 255  
QY 192 -----AONDAGS-----TLRNAEWMKAMGMWGLOARRRAR-DEOVAMQAAASD 235  
DB 256 WRTFRLENNKADNDAGNAIDSLNVEYKYFN-----NEXEADRGDOFLKTYETASL 310  
QY 236 AGGAVMSGIKVFNIVQTLILGGAVLADGKISAGAMTAGSLVGRALAPITGAVGOMK 295  
DB 311 KSTSTLAMNFGONAFISVGLTAIMVLSQG-IVAGALTVGDIVM-----VNGLLFQLS 363  
QY 296 ---NYIG-----ARGA---WDRLOTMLREKSDDHMPLEPPGVLSAEASTILPGAQ- 343  
DB 364 LPLNFGSTVYRETRQALIDMNTLFTLKVTTRKDKYMAPPLQ--ITPQATAYAFDNVHF 421  
QY 344 -----OPTMQASFRIIDAGAAVALVGPSAGKSSLLNGIYVWPCAGAVIRLDGYDIKQW 398  
DB 422 EYIEGQKVLNGSFEVPAAGKVAIVGSSGSKSTIVRLFRFEPOKGSYLAGQNLQDV 481  
QY 399 DPEKGRHVGLPDQLELFSGYVAQNIARFT-EFESQEVTEAATLGVHEMISLPMGYD 457  
DB 482 SLESIRAVGVVPQDAVLFNHTIYINLYGNINASPEEVYAVAKGLADHDAIIRMPGYD 541  
QY 458 TAIGEGASISGGORLALARAVERFPALLVDEPNASLDQGEVALMEAMRLKAAR 517  
DB 542 TOVERGLKLSGGEKQVATARATLKNPVIYLDEATSSLDSTIEETITLAMRV-VKHR 600  
QY 518 TVIFATKVNLLAQAADYIMVINGVISDFG 547  
DB 601 TSIFIAHRLSTVVDADIELVLSGKVAERG 630

RESULT 6  
US-09-195-391-3  
; Sequence 3, Application US/09195391  
; Patent No. 6080842  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Shah, Puryi  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN ATP-BINDING CASSETTE  
; TITLE OF INVENTION: TRANSPORT PROTEIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/195,391  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/895,522  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0336 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:









QY 511 -RLKAKRVIEFATHKVNLLAQAADYIMVINGVISDFGE 548  
Db 621 DNFTSGSRTSVYIAHRLRTIADADKIIVLNGRVEEGK 659

## RESULT 11

US-08-784-649A-2  
; Sequence 2, Application US/08784649A  
; Patent No. 5830697  
; GENERAL INFORMATION:  
; APPLICANT: Sikic, Branimir I  
; APPLICANT: Chen, Gang  
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO  
; TITLE OF INVENTION: CYCLOSPORIN MODULATION  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/784,649A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: Reg No. 5830697 36,677  
; REFERENCE/DOCKET NUMBER: 06037/007001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-322-5070  
; TELEFAX: 415-854-0875  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1279 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-784-649A-2

Query Match 13.5%; Score 379.5; DB 2; Length 1279;  
Best local similarity 24.6%; Pred. No. 2,4e-27;  
Matches 142; Conservative 99; Mismatches 230; Indels 107; Gaps 18;

QY 26 AMVTSFFINILALVSPMLQVDRVLSRVSTL-----IVLVICVFLVYGLLEAL 80  
Db 728 AITPSKITGVFTRI-----DDPEFKRONSLFSLFLALGIIISFTFFLOGFTFGK 778  
QY 81 RTQVLVRCGLKFDGVARDPIKFSVL-----DSTLSRKIGGQ 117  
Db 779 AGEILTKR-----LRYMVFRLRQDVSVFDDPKNTGTALTRLANDAQVGAIGSR 831  
QY 118 AFRMDQVREFMTGGLAFCDAPTPVFAIVSWMLHPFFGILAIACIIIFG---LAVMN 174  
Db 832 LAVITQNLANIGTGIITISF-----IYGMQUTLL--LLAIVPIAIVGVEMKMLS 879  
QY 175 DNATKNPIQAMATMASIAQ--NDAGSTLRNAEYKAKGMWG-GIQARRARRDQVAMQA 231  
Db 880 GQALKKKELEGAGKRIATEALENRTVSLTQEQKFEHMYAQSLQVPR-----928  
QY 232 AASDAGAVMSGIVK-FENIVQTLILG---GAYLAI-----DGKISAGMIAGSIL 279  
Db 929 --NSLRKAHIGITFTFOAMMTFSYACGFRFGAIVLAHKLMSEFDVLLVSAVFGAMA 986

QY 280 VGR--ALAPIEGAVGQKNITGARGAMDRL-----QTMLEEKSSADDHMLPPEPRGLSA 332  
Db 987 VGOVSSFP-----DIAKAKISAHHIMIEKTEPLIDYSTEGMLNTEGNTFG 1037  
QY 333 EAASILPQGAQOPTMROASRIDAGAAVALVGPSAAGKSSLRIGVGVMPACAGVIRLDG 392  
Db 1038 EVVFNYPTRPDIPVLOGLSEVKKGGTLLALVSSGCGKSTVVOLLSEFFYPLACKVILDG 1097  
QY 393 YDIKQWPEKLRHVGYLPDIDIEFSGTVAONIA--RTEFESQEVIEAATLAGVHEMT 449  
Db 1098 KEIKRLNWQMLRAHLGIVSQEPILFDCSIAENITAYGDSRVSQEIVRAAKEANIHAFL 1157  
QY 450 QSLPMGYDTAIGEGASLSGQQRRLAARVFMPLVLDENASLDQVGEAALMEAM 509  
Db 1158 ESIPNKYSTVGVGKGTQSGQKQRIARALVQPHILLDEATSAIDRESEKVVQEAL 1217  
QY 510 RLKAKRVIEFATHKVNLLAQAADYIMVINGVISDFG 547  
Db 1218 DKAREG-RTCIIVAHRLSTIONADLIYVFONGRVKEIG 1254

## RESULT 12

US-08-583-276-19  
; Sequence 19, Application US/08583276  
; Patent No. 5837536  
; GENERAL INFORMATION:  
; APPLICANT: McDonagh, Kevin T.  
; APPLICANT: Nienhuis, Arthur  
; APPLICANT: Tolstoshev, Paul  
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cecchi & Stewart  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: DM4.V2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583,276  
; FILING DATE: 05-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/332,444  
; FILING DATE: 31-OCT-1994  
; APPLICATION NUMBER: 07/887,712  
; FILING DATE: 22-MAY-1992  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1280 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: protein  
; US-08-583-276-19

Query Match 13.5%; Score 379.5; DB 2; Length 1280;  
Best local similarity 24.6%; Pred. No. 2,4e-27;  
Matches 142; Conservative 99; Mismatches 230; Indels 107; Gaps 18;

QY 26 AMVTSFFINILALVSPMLQVDRVLSRVSTL-----IVLVICVFLVYGLLEAL 80  
Db 729 AITPSKITGVFTRI-----DDPEFKRONSLFSLFLALGIIISFTFFLOGFTFGK 779

[illegible]

```

Db      833 LAVIQNIAINLGITGIIISF-----IYQWQTLTL-LLAIYPIIAIAGVEMKMLS 880
OY      175 DNATNPLOMATMASIAAQ--NDAGSTIRNAEVRKAMKMG-GLOARMRAREDOVAMQA 231
Db      881 GQALDKKEELBAGKIAIEAIEENFRTVVSLSIOBKFEEHMYAOSLOVPYR----- 929
OY      232 AASDAGVMSGIXY-FNNIVOTLLLG-----GAYLAT-----DGKISAGMIAIGSL 279
Db      930 --NSLRKAHIGITPISFTQAMMYFSYACCFRGAIVLAHKLMSEFDVLLVSAVVEGMA 987
OY      280 VGR--ALAPIEGAVGQMNNTYIGARGAMDRL-----QTMLEEKSSADHDMPLEPBGVLSA 332
Db      988 VGVVSSFP-----DYAKAKISAHIIIMIEKTLPIIDYSTIEGLMPNLTSENVTFG 1038
OY      333 EAASLLPFGAOPITRKQASFRIDAGAAVALGPSAAGSSLLRIGIVGPCAGVIRLDG 392
Db      1039 EVFENVPRPDIPLQGLSLFEYKKQOTLALVSSCGCKSTVVQLLEFRFDPLAGKVLIDG 1098
OY      393 YDIKWDEKLGRIHGYLPRODELFSGTVQAQIA--RTEFESEGVIEAATLACGVHEMI 449
Db      1099 KEIKLNVQWLEAHIGIYSQEPILFDCSIAENIAYGDSNRVVSOREIVRAKEANIHAFI 1158
OY      450 OSILPGYPTAIEGSGASISGQORLALARAFAFRMPALLVDEPNASIDOVGYLMEAM 509
Db      1159 ESLPRKYSTKVDKQTOISGCGKQKQIALARALVRPHILLDEAISALDTSEKVVQDAL 1218
OY      510 KRLKAKRPTVIFATRKVNLLAODVIMVINGVISDFG 547
Db      1219 DKAREG-RTCIYIAHRLSTIQMADLIVFQNGRVKHEG 1255

RESULT 14
US-08-752-447-2
Sequence 2, Application US/08752447
Patent No. 5994088
GENERAL INFORMATION:
APPLICANT: Mechettner, Eugene
APPLICANT: Koninson, Igor B
TITLE OF INVENTION: Methods and Reagents for Preparing and
TITLE OF INVENTION: Using Immunoligal Agents Specific for P-glycoprotein
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: McDonnell Boehnen Hulbert & Berghoff Ltd.
STREET: 300 South Wacker Drive, Seventh Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752.447
FILING DATE: 15-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5994088nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 95,1121
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-9001
TELEFAX: 312-913-9808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-752-447-2

```

Query Match 13.3%, Score 373.5; DB 2; Length 1280;  
Best Local Similarity 24.6%, Pred. No. 8.9e-27;  
Matches 142; Conservative 96; Mismatches 233; Indels 107; Gaps 18;

OY 26 AMVPSFFINILALVSP-----LYMLQYDRLVLSRNVSTLLVLYVCVLF-----LV 73  
DB 729 AIFSKIIIGVFTRI-----DDPETRKONSLESLFLALGIIISITFFLOGLTFRGK 779  
OY 81 RTQVLRGGLKFDGVARDPFKSVL-----DSTLSRKIGGQ 117  
DB 780 AGEILITR-----LRVFRSMRLRODVSMPHDPKNTGALTTLRLANDAOVKALISGR 832  
OY 118 AFRDMQVREEMTGLLAFCDAPWTPVEVYSWMLHPFGLIATIIIGLAVMN--- 174  
DB 833 LAVITQNIANTLGTIIISF-----IYGMQLTL--LLAIYPIIATAGVEMKKEFA 880  
OY 175 DNATNPLOMATMASIAAO--NDAGSTLRNMAEVMKAMGMWG--GLOARRARDEQVAMQA 231  
DB 881 GOALDKKKELEGAGKIATDEAIENFRFVSVSLTQEQKEHHYVQASLOVPYR----- 929  
OY 232 AASDAGAVMSGIKV--FRNIYQTLILG---GAYLAI-----DKRISAGAMTAGSIL 279  
DB 930 --NSLRKMHIEGTFSPFOGAMWFSYAGGFRGATLVAKHKLMSFEDVLLVFSAVVFGAMA 987  
OY 280 VGR--ALAPLEGAVGOMKNYIGARGAMDRL-----QTMIREKSSADHMPDPERGULSA 332  
DB 988 VQGVSSFP-----DYAKAKISAHHIIMIEKTPIDBSYSTEGLMPNTLEGVTFEG 1038  
OY 333 EASLILPPGAQOPTMRQASFRIDAGAAVALVGPASAGKSSILRGIVGWPCAGVIRIDG 392  
DB 1039 EYVENYPRPDIPVLQGLSLEVKKQOTLALVSSGCGKSTVVQLERFDPDLAGVLLDG 1098  
OY 333 YDIKQDEPEKLGRRHVGYPQDIELEFSGTVAONIA--RTEFESQEVIEAATLAVGHEMI 449  
DB 1099 KEIKRLANQWMLRAHGIYSQEPILFDCAIAENIAGDNSRVVSGEIVYRAAEANIHAFI 1158  
OY 450 OSLPMGYDTAIGEGASISGGORQLALARAVERMPALIVDEPNASLDQVEVALMEAM 509  
DB 1159 ESLPKYSKRVKDGKTOISGGOKORIALRRLVROPHILLDEATSSALDTSEKVVQOAL 1218  
OY 510 KRLKAARTVIEFATKRVNLLAQADYIMVINOVISDFG 547  
DB 1219 DKAREG-RICIVIAHRLSTIONADLIVFQNGRVKEHG 1255

RESULT 15  
US-08-232-537-2  
Sequence 2, Application US/08232537  
Patent No. 5516655  
GENERAL INFORMATION:  
APPLICANT: Peery, Robert B.  
APPLICANT: Skatrud, Paul L.  
TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF  
TITLE OF INVENTION: AUREOBASIDIUM PULLULANS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,537  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Plant, Thomas G.

REGISTRATION NUMBER: 35,784  
REFERENCE/DOCKET NUMBER: X9212  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-2459  
TELEFAX: 317-276-1917  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1302 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-232-537-2

Query Match 13.3%, Score 373.5; DB 1; Length 1302;  
Best Local Similarity 25.1%, Pred. No. 9.1e-27;  
Matches 151; Conservative 106; Mismatches 205; Indels 139; Gaps 26;

OY 26 AMVPSFFINIL-ALVSP-----LYMLQYDRLVLSRNVSTLLVLYVCVLF-----LV 73  
DB 740 AIVEGNLISVLNDESPDRSRADLFSLF-----ILALIALEFSYAGNGCC 786  
OY 74 YGL-----LEALRTQVLRGGLK--EDGVARDPFKSVLSDSTLSRKIGGQAFRDMQ 124  
DB 787 FGIVSHFVAKIOHISLASILRQDMQFSGOSVPSLMSLSDAGOLACISVAIGTIFT 846  
OY 125 VREFMTGCLIAFCDAWTPVEVYSWMLHPFGLIATIIIGLAVMNDATKNPLOM 184  
DB 847 VCVSITGGII-----LAHVVAWK-----IAVLLAAVPVW---ITAGVYRL 884  
OY 185 ATMASI-----AAONDAOS-----TRNMEVVKAMGMGLOARRARDEQVAMQAASD 235  
DB 885 RVLALAESRHRSAVYNDASIAEACRGIRTIASIGREGV-----SRASN 929  
OY 226 AGGAVM-----SGIKVERNIYQTLI-----LGGAYLAIDGKISA 270  
DB 930 A-AVKEPYDKIR-FTLITNTLLALSFITYFVALAYWMCARKYRNGTISQDFFIVL 986  
OY 271 GAMINGSLVGR--ALAPLEGAVGOMKNYIGAR--GAMDRLOTMLREKSSADHMP--- 322  
DB 987 PALFSAQSAGQIFSLSPMSRAG-----VAARNVFGIHDQKPTIIVDAKOSGALPESST 1041  
OY 323 -----LPEPRGVLSAEASIT-LPPGAQOPTMRQASFRIDAGAAVALVGPASAGKS 371  
DB 1042 LSIPLEDKASPSGSGWIEFKNVSLCYPSKQHPALQVNNISIRGERFIALVPSGAGKS 1101  
OY 372 SLIRGIVGWPCAGVIRIDGYDIKQMD--PEKIGRRHVGYPQDIELEFSGTVAONI---AR 427  
DB 1102 TILSLQRRYDPTAGSVQDGDIREVAVPQHRGR-LGIYPOEPDLFGSISYNTGLGA 1160  
OY 428 FTEPESQEVIE-AATLAGVHEMIQSLPMGYDTAIGEGASISGGORQLALARAVERMPA 486  
DB 1161 POLQVTRDIDIEKICAKGHEIFIMSLPEGYSTECGNSKSLGGOKORIAVARALIRPSE 1220  
OY 487 IIVDEPNASLDQVEVALMEAMKRLKAARTVIEFATKRVNLLAQADYIMVINOVISDF 546  
DB 1221 VLLDEYTSALDAHQEQITKAVDG-ASVDRITIVVAHRLSTIONADRLVFDGRVVEV 1279  
OY 547 G 547  
DB 1280 G 1280

Search completed: July 19, 2002, 10:41:36  
Job time: 36 sec



THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:41:00 ; Search time 32.83 seconds  
(without alignments)  
1624.416 Million cell updates/sec

Title: US-09-913-414-4

Perfect score: 2807  
Sequence: 1 MFRSGAKPTIFDQAVLVAR.....MNINGVISDFGERDRCPWS 555

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2765	98.5	578	D87374	RsaA secretion sys
2	1231	43.9	593	S26696	alkaline proteins
3	1136	41.2	600	C83221	transport proteins
4	1077	38.4	582	D96003	ABC transporter AT
5	1065.5	38.0	602	AF0477	ABC transporter pr
6	1044.5	37.2	564	E70394	ABC transporter (h
7	1039	37.0	576	A49933	proteinase SM tran
8	1039	37.0	585	T48672	ABC-type transport
9	1029	36.7	615	E96310	rhilobidin secret
10	1029	36.7	615	AF2972	alkaline proteins
11	1026.5	36.6	575	S12525	metalloproteinase
12	1012	36.1	583	E71687	hypothetical prote
13	1011	36.0	584	E97753	hypothetical prote
14	958	34.1	589	D95953	hypothetical prote
15	935	33.3	577	AH2908	hypothetical prote
16	935	33.3	588	G97683	prsd protein (U891
17	931	33.2	690	D96318	ABC protein AGR_L
18	930	33.1	583	A13098	hypothetical prote
19	930	33.1	583	H98187	prsd protein (U891
20	927	33.0	551	A12964	hypothetical prote
21	609	21.7	989	AE2140	hypothetical prote
22	607.5	21.6	1003	AE2335	toxin secretion AB
23	592.5	21.1	706	T00228	hemolysin B - Esch
24	584.5	20.8	707	A61378	leukotoxin express
25	577.5	20.6	1011	S75806	hemolysin secretio
26	576.5	20.5	708	C30169	leukotoxin express
27	575.5	20.5	725	T03514	probable secretion
28	573.5	20.4	975	AC2517	hypothetical prote
29	572.5	20.4	742	H81855	probable cytolysin

30	569.5	20.3	707	2	D43599	hemolysin I secret
31	567.5	20.2	742	2	H81088	ABC transporter fa
32	562.5	20.0	707	1	LEEBBV	hemolysin secretio
33	562.5	20.0	720	2	E82561	toxin secretion AB
34	561.5	20.0	707	2	S10057	ABC-type transport
35	559.5	19.9	707	1	LEECB	hemolysin secretio
36	559	19.9	893	2	AH2007	toxin secretion AB
37	552.5	19.7	711	2	C49219	toxin apxIII secre
38	549.5	19.6	723	2	C83412	probable ATP-bindi
39	544	19.4	720	2	H82198	RTX toxin transport
40	543.5	19.4	704	2	H82381	toxin secretion AT
41	533	19.0	712	1	BVBRCB	cyab protein - Bor
42	531.5	18.9	706	2	G90696	hypothetical prote
43	531.5	18.9	720	2	C85547	probable cytoplasm
44	522	18.6	712	2	AD0274	probable toxin tra
45	521.5	18.6	718	2	AF0835	probable type I se

ALIGNMENTS

RESULT	1	
D87374		
RsaA secretion system, ATP-binding protein RsaD [imported] - Caulobacter crescentus		
C:Species: Caulobacter crescentus		
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #extl_change 20-Apr-2001		
C:Accession: D87374		
R:NIEMAN, W.C.; FELDBLYUM, T.V.; PAULSEN, I.T.; NELSON, K.E.; EISEN, J.; HELDEBERG, B.; LAUB, M.T.; DEBOY, R.T.; DODSON, R.J.; DURKIN, A.S.; GILIN, M.L.; HAFT, D.H.; KO N, J.; ERMOLEVA, M.; WHITE, O.; SALZBERG, S.L.; SHAPIRO, L.; VENTER, J.C.; FRASER, C Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001		
A:Title: Complete Genome Sequence of Caulobacter crescentus.		
A:Reference number: A87249; MIMID:21173698; PMID:11259647		
A:Accession: D87374		
A>Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-578 <STO>		
A:Cross-references: GB:AE005673; NID:g13422298; PIDN:AAK22992.1; GSPDB:GM00148		
C:Genetics:		
A:Gene: CC1008		
Query Match	98.5%	Score 2765; DB 2; Length 578;
Best Local Similarity	99.8%	Pred. No. 9.1e-189;
Matches 549; Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
QY	1	MFRSGAKPTIFDQAVLVARPAVITAMVSEFTNIALVSPLYMQLQYDRLTSRNVSTL 60
DB	1	MFRSGAKPTIFDQAVLVARPAVITAMVSEFTNIALVSPLYMQLQYDRLTSRNVSTL 60
QY	61	IVTIVTCVPLFLVYGLLEALRTQVLYRGGLKPDGVARDPIFKSVLDSTLSRKIGGOARR 120
DB	61	IVTIVTCVPLFLVYGLLEALRTQVLYRGGLKPDGVARDPIFKSVLDSTLSRKIGGOARR 120
QY	121	DMQOVEFMFGILAFCDAPWTPFVIVSMILHPFGILAIICIIIFGLAVNNDATKN 180
DB	121	DMQOVEFMFGILAFCDAPWTPFVIVSMILHPFGILAIICIIIFGLAVNNDATKN 180
QY	181	PIQMATWASIAQNDAGSTLRNAEYVKAKGMWGGILQARRMARREDEQVAMQAAASDAGAV 240
DB	181	PIQMATWASIAQNDAGSTLRNAEYVKAKGMWGGILQARRMARREDEQVAMQAAASDAGAV 240
QY	241	MSGIKVFNNIVQTLILGGGAYLAIDKISAGAMIASILVGRALADIEAGVGMKMYICA 300
DB	241	MSGIKVFNNIVQTLILGGGAYLAIDKISAGAMIASILVGRALADIEAGVGMKMYICA 300
QY	301	RGAMDRLQTMLEREKSAODHMLPEPRGVLSAFAASILPPGAQOPTMROASPRIDGAAY 360
DB	301	RGAMDRLQTMLEREKSAODHMLPEPRGVLSAFAASILPPGAQOPTMROASPRIDGAAY 360
QY	361	ALVGPAAKSSILRGIIVGVWPCACAVIRLDGYDIRKQMPDKLGRHVGYLPDIDELFSGT 420
DB	361	ALVGPAAKSSILRGIIVGVWPCACAVIRLDGYDIRKQMPDKLGRHVGYLPDIDELFSGT 420

QY	421	VAONLARFEEFSQVIEATL	AGVHEMIQSPMGYDPA	IGGGSLSGGQ	QORQALARA	480
Db	421	VAONLARFEEFSQVIEATL	AGVHEMIQSPMGYDPA	IGGGSLSGGQ	QORQALARA	480
QY	481	VERPALLVLDEPNASLDO	VEVALMEAKRKLRKAK	RTVTFATHKVN	LAAQADYIMVINO	540
Db	481	VERPALLVLDEPNASLDO	VEVALMEAKRKLRKAK	RTVTFATHKVN	LAAQADYIMVINO	540
QY	541	GVISDGERD	550			
Db	541	GVISDGERD	550			

## RESULT 2

alkaline proteinase secretion protein Aprd PA1246 [imported] - Pseudomonas aeruginosa  
C.Species: Pseudomonas aeruginosa  
C.Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 02-Feb-2001  
C.Accession: S26696; F83489  
R.Duong, F.; Lazdunski, A.; Caml, B.; Murgier, M.  
Gene 121, 47-54, 1992  
A.Title: Sequence of a cluster of genes controlling synthesis and secretion of alkaline  
A.Reference number: S26696; MUID:93051361  
A.Accession: S26696  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-593 <D>  
R.Stover, C.K.; Pham, X.B.; Ervin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yvan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim  
..; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A.Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho  
A.Reference number: AB2950; MUID:20437337  
A.Accession: F83489  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-593 <S>  
A.Cross-references: GB:AE004554; GB:AE004091; NID:93947174; PIDN:AA604635.1; GSPDB:GN001  
A.Experimental source: strain PAO1  
C.Genetics:  
A.Gene: aprD; PA1246  
C.Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
C.Keywords: ATP, nucleotide binding; P-loop  
E.349-343/Domain: ATP-binding cassette homology <ABC>  
E.366-373/Region: nucleotide-binding motif A (P-loop)

Query Match	43.9%	Score 1231	DB 2	Length 593
Best Local Similarity	46.5%	Pred. No. 9.2e-80		
Matches 251	Conservative 109	Mismatches 178	Indels 2	Gaps 2

```

0Y 14AVLVAAPVATITAMVFEFFINIIATLSPILMLOYDDRLTGRNASTLVIVPVCLEFLV 73
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 14QALAAARGALRVAABEGVINLLMLVPSILMLOYDRILSSANEVTLMLTLMLGVEYF 73
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
0Y 74YGLLEALRTQVLVGRGKLEFDGVARDPIFKSVLDSTLSRKG-IGGAFRMDQVREFPMFG- 131
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 74MGALTEALRSVLVAVRSEFPGQLHGRITVAAFAFERNLIRAGGQEAQSALHDLTFLRQFITGQ 133
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
0Y 132GLIAFDCAPEPTPVFVITSMWLIHPEFGILATIIACIIIEGLAMINDATKNPIOMATMASIA 191
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 134ALFAFPDAPFPPVLLVITFEDPWLGLSTLVGALALMALAMFNERATRAPLAKAABELSTIK 193
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
0Y 192AQNDAAGSTLLENAEYMKAMGMWGLQABRRARDEQVAAQAAASDAGAVMSGIKYFERNIV 251
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 194SGQLASNNLENAEYIEAMGMLGSRKMERBRLHQAFLLDQGLASEARAIRINLALSTYLRIAL 253
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
0Y 252QTIILGGATVLADGKISACAMITGSLIVGNALAPIBEGAVGOMKNTIARGAMDRLQTM 311
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 254QSVLVIGAGVLAVEGRITTPGMMITGSLIMGRALPIDIOLIGWKMGKGAARDAYRRIRISGLL 313
   ||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

```

QY	312	REESADDDHMLPEERGRVLSAEASLITPRGMOQPMRQASRIDAAGVAVLVGSPAAKS	371
Db	314	DEFPARERRMELPERGRHLLLESLDAPBPGEARTLRLTALIPGVSVYIGBGSCKS	373
QY	372	SLLRIRVGVMPFACAVIRLIDGYDIKOMPPEKIRGVGLIPDILIEFSGTVQANTARTEF	431
Db	374	SLARVVLGWPTLHGSVSLDGAELIRQYERETPLGRITGLPDILIEFACTVANTAREEV	433
QY	432	ESQEVTEAATLGVHEMIQSLPMTGYDTAIGBGSLSGSGORLALANAVRPMALLVD	491
Db	434	QADKVVVEAARLGVHEHLRLPFGYDTVLGVGAGALSOGORIRALALVGAPLVLLD	493
QY	492	ENNASLDDVGEVLEALMEAKRRLKAARKTVIPATHKVNLLAADYIVAVINQYISDGEEDR	551
Db	494	EVNSNLDISGEGALLAALQALKARCTVLLITTHRGVLSGCDRLRLALNAGQILHLYGEDQ	553

### RESULT 3

transport protein Hasd PA3406 [imported] Pseudomonas aeruginosa (strain PA01)  
C.Species: Pseudomonas aeruginosa  
C.Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C.Accession: C83221  
R.Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; L  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L  
: Lorry, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A.Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A.Reference number: AB2950; MUID:20437337  
A.Accession: C83221  
A.Status: Preliminary  
A.Molecule type: DNA  
A.Residues: 1-600 <STO>  
A.Cross-references: GB:AE004761; GB:AE004091; NID:99949533; PIDN:AA606794.1; GSPDB:GN  
A.Experimental source: strain PA01  
C.Genetics:  
A.Gene: PA3406  
C.Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match	41.2%;	Score 1156;	DB 2;	Length 600;
Best Local Similarity	43.48%;	Pred. No. 2e-74;		
Matches 230;	Conservative 117;	Mismatches 181;	Indels 2;	Gaps 2;

```

Oy 20 PRAVITAMESEFINILAVSPYMLQYDVRVTSRNNSTVLVEVIFVFLVGLLEA 79
    ||| : : : ||| ||| ||| ||| ||| : : : ||| : : : ||| :
Db 20 PRAKLSVALEFVAVINILMLAPSLYMLQYDRLVSGNNHMTLMLTLMVLGLYLLGALAEW 79

Oy 80 LFTOVLVRGKLFKDVARDPIFKSVLDSTLSR-KGIGQAFRDMQDVEEPMTG-GLIAFC 137
    : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 VRSVLVIRLGGGLDMQJLNDQNRIDYDASFRASTLEGEQAGALNDLTLSRQFLTGNALPAFF 139

Oy 138 DAPMTPVAFIVASWMLHPFGILATIACIIIFELAYMNDNATNPNLOMTMMSIAQNDAG 197
    ||| : : : : : : : : : : : : : : : : : : : : : : :
Db 140 DAPWPEPLYLVLFELTSPWGLTLLAGALLLVLLAVNERSRSEPLAEAGQUSTIATQGAS 199

Oy 198 STLRLAEVVKKAGMMNGGLOARWRARREDECVAMQAAASDAGVMSGIKYFRNIVOTLLIG 257
    : : : : : : : : : : : : : : : : : : : : : : : : :
Db 200 AULRQAEITLAAAGMLPAMARARFADQALFLAQNIGSESAIIGATSKGVRALDQSLVIG 259

Oy 258 GGAYTALIDKISAGAMIAGSILVGRALADIEGAVGOMKNYITGARGAMDRLQTMLEEKSA 317
    ||| : : : : : ||| ||| : : : : : : : : : : : : : : :
Db 260 LGMALVADDLITPGMMIAGSILMGRLSPIDQLIAMVRQMSGARQAYORLARLTLEENPA 319

Oy 318 DDHMTLPBERGVLSEAFASILPPGAQOPMROASFRIDAGAVALVGPSPAARKSSLLGI 377
    ||| ||| ||| : : : : : : : : : : : : : : : : : : : :
Db 320 ALGMPLPARAGALRVERLICAAAPGRQALLDOLGFALEGEALVIGIPGSGKSTLARL 379

Oy 378 VGVMPPLACAVIRLDGDKIQOMDPEKIGRHVGLIPDDIEFSGTVQONIARTEFESEVYI 437
    || : : : : : ||| ||| : : : ||| ||| : : : ||| : : : ||| :
Db 380 VGMQPLSGAVVLDDGDLRQMSAALGPHITIGLADQVOLFASISIENTARFVEVAEKVY 439

Oy 438 EATLAVGHEMTQSUPMGYDTAIGEGGASLSCGORORLALARAVERMPALLVDEPNASL 497

```



Db 440 AAARLGVHDLVRLPGGYDTRLGDGAGLGGGGRIGRLALYGRPALIVDEPNASL 499  
QY 498 DOYGEVALMEAMKRLKAARTVIFATHKVNLAAQADYIMVINGVISDFG 547  
Db 500 DEAGEAALAEAIAMRRGSSLVLVTHKPAVLALTDKLLHGGRLQRF 549

## RESULT 4

D96003

ABC transporter ATP-binding protein prsd [imported] - *Stenotrophobium meliloti* (strain 102)  
C:Species: *Stenotrophobium meliloti*  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: D96003  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb *psymb* megaplasmid from the N2-fixing endo  
A:Reference number: A95842; M0ID:21396508; PMID:11481431  
A:Accession: D96003  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-582 <KUR>  
A:Cross-references: GB:AL591985; PIDN:CAQ49692.1; PID:g15141179; GSPDB:GN00167  
A:Experimental source: Strain 1021, megaplasmid *psymb*  
R:Gallibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
heault, F.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont *Stenotrophobium meliloti*.  
A:Reference number: A96039; M0ID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: prsd. Smb21466  
A:genome: plasmid

Query Match 38.4%; Score 1077; DB 2: Length 582;  
Best local Similarity 42.2%; Pred. No. 7,9e-69;

Matches 226; Conservative 104; Mismatches 202; Indels 4; Gaps 3;

QY 20 RPAVITAMVFSFPIITALVSPMLQYDVRVLTSRNVSTLIVTVYCVLFLVYGLLEA 79  
Db 19 RAAFIGVAVASALVNLTYLTGSEFMLEVDRLIPRSRISPLIALSLALLLVAFQAFEL 78  
QY 80 LRQVYLRGGLKDVARDPIFKSVLDSTLS--RKGIIGQAFDMQDQVREFMTG-GLIAR 136  
Db 79 IRGRMLVRIAGALDESLNGRIYRAIVKAPLKLMOGDGLQALRDPQVRSFSLGSGFPAAL 138  
QY 137 CDAPWTFVFIYVSMMLHPFEGILAITACITIFGLAVMNDATKNTPIOMATMASIAQONDA 196  
Db 139 FDLPRMPLFYAICFLFRHYVIGLAIIGGLITLITLTNNGTQAPAKKASEAGCLRNVFA 198  
QY 197 GSTLRNAEYKAMGWMGGLQARWRARDEQVAAQAAADAGAVMSGKIFRNIIVOTLIL 256  
Db 199 OASORNAEYVHAAMGSARLTALTEWRNTERFERDNRTSDINGYGALSKYFRALQSGVL 258  
QY 257 GGAAYLAIQKISAGAMIASIIVGRALPIEGAVGQWKNTICARGAMRLQTMLEPKS 316  
Db 259 AAGAVLVIRKESAPGIIIGSITLARALPVELAIGNMRLVNAAROSWOLKELNALPE 318  
QY 317 ADQHMPLPEPRGVLSAEASILPRGAOOPTRQASFRIDAGAVLVGPSAACKSSILRG 376  
Db 319 ADAPRLQPRPHEKTLVEGLASGRPAAROLVSVNFTVRAAGAVGYIGSASAKSSILARA 378  
QY 377 YGVWPCAAGVLRDGYDIKQMDPEKLRHNVGLPQDIELFSGTVAQNTARFTE-ESQ 435  
Db 379 ILGIWPAVSGSVRLDGAALDQMSDALGKHVGYLPQDVELFACTIQAONICRFADATSEA 438  
QY 436 VIEAATLAGVHEHIOSLPMGYDTAIGEGASLSGGORQRLALARAVERFALLVDEPNA 495  
Db 439 IVAAAAARVNDILRLPNGYDTEIGDGWTLISAGORORVALARALYGDPELVVLEPN 498

QY 496 SLDOVEGVALMEAMKRLKAARTVIFATHKVNLAAQADYIMVINGVISDEGRD 551  
Db 499 NLDAEGQALSEALMSVRSRGIVYVAHRPSALASDVLMMEGMRQAFPRQ 554

## RESULT 5

AF0477

ABC transporter protein ypo3920 [imported] - *Yersinia pestis* (strain C092)  
C:Species: *Yersinia pestis*  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C:Accession: AF0477  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M  
deno-Tarrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davies, P.; Dougan, G  
llo, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A:Reference number: AB0001; M0ID:21470413; PMID:11586360  
A:Accession: AF0477  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-602 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAQ93386.1; PID:g15981832; GSPDB:GN00175  
C:Genetics:  
A:Gene: YPO3920

Query Match 38.0%; Score 1065.5; DB 2: Length 602;  
Best local Similarity 41.5%; Pred. No. 5,4e-68;

Matches 227; Conservative 108; Mismatches 203; Indels 9; Gaps 5;

QY 5 SGAKPTTFDOAVVAVARAVITAM-VFSFPIITALVSPMLQYDVRVLTSRNVSTIYL 63  
Db 10 SSVPPTIL--SVLAGNKKIIMGIGLFTAVNLMLAPAIMLOVYDRLVASANTMTLML 67  
QY 64 TVICVFLVYGLLEALRTQVLYRGGLKFDGVARDPFKSVLDSTLSRK--IGGQAFR 120  
Db 68 TVLVAGVFVFIIGLEWVRSVAVIRLGTQIDMQNPENAFNAANL--KGNHNPAAQALN 125  
QY 121 DMOQVREFMTG-GLIARCDAPWTFVFIYVSMMLHPFEGILAITACITIFGLAVMNDATK 179  
Db 126 DLTVLRQFATGNALFAFFDAPWPEFLVLYFLFLPMWGLMAAGAGLLVYLAAMLNQWICK 185  
QY 180 NPIOMATMASIAQONDAQSTLRNAEYKAMGWMGGLQARWRARDEQVAAQAAADAGGA 239  
Db 186 KPLHDAISITSHATQOANANLRNADVIEMAGMLKALREMLMOHANFLYQANLASDKSSR 245  
QY 240 VMSGKIFRNIIVOTLILGGAYLAIQKISAGAMIASIIVGRALPIEGAVGQWKNTYIG 299  
Db 246 VTAIVAKSSRQALQSMMLGGLALLVYNEIRTAGVMIAGSILIGRYLQIDQLIYVWKQMSH 305  
QY 300 ARGAMDRLOTMLEPKESADDMPLPEPRGVLSAEASILPRGAOOPTRQASFRIDAGAA 359  
Db 306 ARLAYORLSQALQHPSPGFWLPAPOGKLVNTOLMACRPGTHIPVLSINFELOPGDV 365  
QY 360 VALVGPSAACKSSILRQIVGWPCAGVLRDGYDIKQMDPEKLRHNVGLPQDIELFSG 419  
Db 366 LGTIGFSGSKTLAKULVASOPTFSGTVLDSADLSRMDKTQDGERIGTLPQNTQIFRG 425  
QY 420 TVANQNTARFTEFSOEYEAATLAGVHEHIOSLPMGYDTAIGEGASLSGGORQRLALAR 479  
Db 426 TVENATARFGAIDTAKVNAAGLADVHDLTLHPGQYDPLDGDGEGSLSGGQRRLALAR 485  
QY 480 AYVRMPALVLDEPNASLDOVEGVALMEAMKRLKAARTVIFATHKVNLAAQADYIMVIN 539  
Db 486 AMYGIPLVLDEPNASLDEKQALASITQIKQOGCTIVMTIHRBELLSGSDYDLFLK 545  
QY 540 OGVIQSF 546  
Db 546 NGQMDLP 552

RESULT 6

E70394  
ABC transporter (hlyB subfamily) - Aquifex aeolicus  
A:Title: Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 02-Feb-2001  
C:Accession: E70394  
R:Beckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:9819666  
A:Accession: E70394  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-564 <AOF>  
A:Cross-references: GB:AE000723; NID:g2983569; PIDN:AAC07149.1; PID:g2983577; GB:AE00065  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: abcT5  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
C:Keywords: ATP; nucleotide binding; P-loop  
F:349-543/Domain: ATP-binding cassette homology <ABC>  
F:366-373/Region: nucleotide-binding motif A (P-loop)

Query Match 37.2%; Score 1044.5; DB 2; Length 564;  
Best Local Similarity 39.7%; Pred. No. 1.5e-66;  
Matches 212; Conservative 116; Mismatches 203; Indels 3; Gaps 3;  
QY 20 RPAVITAMVPSFINITIALVSPMLQYDRVLTSRNVSTLIVTVICFELVYGLLEA 79  
DB 19 KKTLLIVGLFSLFNLITFLPSIYMLAVDLPVSTPSTLLVITLAVLYFALGLQS 78  
QY 80 LRTQVLRGKFPDGVARDPIFKSVLDSTLSKRC-IGGAFRMDQVREFMTGC-LIAFC 137  
DB 79 VRAVVMQIISLKDESELNKEVTSFEVAIRNPKASQPIINDLYQLKQFLTSVLEFAIF 138  
QY 138 DAPPTPVFVIVSMMLHPEFGILAIITACIIIFGLAVMNDNATKPIOMATMASIAONDAG 197  
DB 139 DLPWPIYFVGLFVHFHYGVMAIISMAVIALILNEYITKKKIKESNELVSTNFLN 198  
QY 198 STLNAAEVMKMGWGCGOARWRARDEQVAMQAAASDAGAVSGIKVFENIYQTLIG 257  
DB 199 RALNAAEVEVLDGMRNNLYKRWNNFYSHLSAFEBADRNFLSLTRIFIMOSLMLG 258  
QY 258 GGAVLAIIDKISAGAMIAGSILVGRALAPIEGAVGOMKNYIGANGAMDRLQTLREK-S 316  
DB 259 LGVLAIIKHETITGVIVAGSILIGRIDPTIYNGWRQIINTKAVIYRLNEPLKFLFK 318  
QY 317 ADDHNPDPERPGVLSAEASILPFGAQOPTMKOASFRIDAGAAVALVGPSAAGSSLLRG 376  
DB 319 REVSVKLPPEKGEIETLSNVVVVPEEGKTPVLRNINMRILPGEFVALIGPSSGSKSLVRT 378  
QY 377 IVGWPCAAGYAIRIDGYDIKOWDEPKGRHVGYPDIELEFSGTVNAONIARFTEESQEV 436  
DB 379 ILGIMLPVHGIVEIDGALDKOMDRYFGKIVGYLPDITELLEGVIAENIARFGEIDSKI 438  
QY 437 IEAATLAGVHEMIOSLPWGYDTAIGEGASISGQORLALARAFAVRPALLVIDEPNAS 496  
DB 439 IEAKLSAHVYIIKLPDGYDTYIGPGITLSGQORIALARALYGNPRVYIIDEPDSN 498  
QY 497 LDQGEVALMEAMKRLKAKRTVIFATHKVNLAQADYIVINGVISDFGERD 550  
DB 499 LDQGEQALYNALIELKRRKVTYIIVSHIRLLMLVDKIAIMODGTLEKAFKAD 552

RESULT 7  
A49933  
proteinhase SM transporter Prtd - Serratia marcescens  
N:Alternate names: inner membrane metalloprotease transporter Prtd  
C:Species: Serratia marcescens  
C:Date: 13-Sep-1994 #sequence\_revision 18-Nov-1994 #text\_change 02-Feb-2001  
C:Accession: A49933  
R:Letoffe, S.; Ghigo, J.M.; Wandersman, C.

J. Bacteriol. 175, 7321-7328, 1993  
A:Title: Identification of two components of the Serratia marcescens metalloprotease  
A:Reference number: A49933; MUID:94042907  
A:Contents: 365  
A:Accession: A49933  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-576 <LET>  
A:Cross-references: GB:s67013; NID:g452927; PIDN:AAB28772.1; PID:g452928  
A:Note: sequence extracted from NCBI backbone (NCBIN:139861, NCBI:P:139863)  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
C:Keywords: ATP; nucleotide binding; P-loop  
F:339-523/Domain: ATP-binding cassette homology <ABC>  
F:356-363/Region: nucleotide-binding motif A (P-loop)

Query Match 37.0%; Score 1039; DB 2; Length 576;  
Best Local Similarity 42.1%; Pred. No. 3.9e-66;  
Matches 222; Conservative 96; Mismatches 195; Indels 14; Gaps 5;  
QY 28 VESFINILALVSPMLQYDRVLTSRNVSTLIVTVICFELVYGLLEALRTQYLV 87  
DB 16 LFTAVINLMIAAPALYMLQYDRVLPBGNKRTILAMILTMVYGLYFGLLEWRSQYVIR 75  
QY 88 GLKFDGVARDPIFKSVLDSTL-SRKIGGQAFRMDQVREFMTG-GLIAPCDAPWTPVF 145  
DB 76 LGAQMDRLNQRVYDAFETMLKGNPLAGQALNDLTLNQFATGNALFAFFDAPMPVY 135  
QY 146 VIVSMMLHPEFGILAIITACIIIFGLAVMNDNATKPIOMATMASIAONDAG--STLRNA 203  
DB 136 LTVFLLHPWIGLALASAGCDRAGAAGLAQSAQCKRLSAERPGSCAHRAGOPANLRNA 195  
QY 204 EVMKMGWGCGOARWRARDEQVAMQAAASDAGAVSGIKVFENIYQTLILGGAYLA 263  
DB 196 EAIAMGMLDRLRWRIRQHQFLLQNRASEKIAAYTAMSKYVRLAQSLMGCGALLA 255  
QY 264 IDKISAGAMIAGSILVGRALAPIEGAVGOMKNYIGARGANDLQTLREKSDHMP 323  
DB 256 VSGDITFGMIAIGSILIGRVLGPIIDOLIGAMKQSSAROSLQREVMANPRIPSLP 315  
QY 324 PEPBGVLSAEASILPFGAQOPTMKOASFRIDAGAAVALVGPSAAGSSLLRGIVGWPC 383  
DB 316 PARGALIVSOLITASAPGTAIPVYHGVSFRUEAEVGLVIGASSGTTLLMRQVLGLTP 375  
QY 384 AAGYIRLDGYDIKOWDEPKGRHVGYPDIELEFSGTVNAONIARFTEESQEVIAATLA 443  
DB 376 ISG-----DGAQGLGPHIGYLPDITQLPAGTITDNIARFGQVDAKVVAAALA 425  
QY 444 GVHEMIOSLPWGYDTAIGEGASISGQORLALARAFAVRPALLVIDEPNASIDQGEV 503  
DB 426 GVHDLIHLPGYETELGEGSGISGQORALARALYGPSALVYIDEPNANDREGEE 485  
QY 504 ALMEAMKRLKAKRTVIFATHKVNLAQADYIVINGVISDFGERD 550  
DB 486 ALQRAIEALKARGNITVLYTHKPAIATTDKLVILVIAQOVHGFPSD 532

RESULT 8  
T48672  
ABC-type transport system ATP-binding protein hasd (validated) - Serratia marcescens  
N:Alternate names: hasa export system ABC-type transport component hasd  
N:Contains: adenosinetriphosphatase (EC 3.6.1.3)  
C:Species: Serratia marcescens  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Sep-2000  
C:Accession: T48672  
R:Letoffe, S.; Ghigo, J.M.; Wandersman, C.  
Proc. Natl. Acad. Sci. U.S.A. 91, 9876-9880, 1994  
A:Title: Iron acquisition from heme and hemoglobin by a Serratia marcescens extracell  
A:Reference number: 224524; MUID:95024064  
A:Accession: T48672  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-586 <LET>

[illegible]

A;Map position: linear chromosome

Query Match	36.7%;	Score 1029;	DB 2;	Length 615;
Best Local Similarity	39.2%;	Pred. No. 2.2e-65;		
Matches 213;	Conservative 118;	Mismatches 196;	Indels 16;	Gaps 5;

```

0Y 22 AVITMVSSEF-----INILALVSPYLMQYDRLVTSRANSTLVLVVICVFLV 73
Db 17 AVVSPKSTFVLGVALVSCVNNIALTSPLEPMQYDYDRLASGSLPTLVGLALLGLXCF 76
QY 74 YGLEALRTOVLVNRGLKFD---GVARDPIFKSVLSDTSLSRKIGGQAFRMDQVREFM 129
Db 77 OCLLDIIIRAVLIRIGEDFDMRYSGRVHADAVVRLPLVNRM--PDGGLOPLRDLNVGFL 134
QY 130 TG-GLIACDAPWTPVPEVIVSMTLHPFGILAIACIIIFGLAVMNDNATNPLOMATMA 188
Db 135 SGTGTAEFFDLPMPELYLGLCFLEFHWIGVTAAGAVLLVSLTITNIFSQKPIRDTWE 194
QY 189 SIAAONDAGSTLRNAAEVKRAMGMMGGLOARRRARREDEVAAQAAASDAGVMSGIFRVER 248
Db 195 NMARRKQLEASRRNAEEVQAMGLGRRLCKRQSSNEATILANRRKAGVAGGLMLASLR 254
QY 249 NIVOTLLIGGAYLAIDKISAGAMIAGSLVGRALAPIEGAVGOMKNYIGARGAMDRLQ 308
Db 255 VVLQAILAVGAMVLIYREQASGVMIASSIMMGRALPVDLAISMKSFEVAQRSMARLR 314
QY 309 TMLREKSAADHMPLEPERGYLSAEASILPPGAQOPLMRQASRIDAGAVALVGPSAA 368
Db 315 ELFAQMPANNAAMALPKPEKELRVENVTIYPGEEKPVTGLGVVTAAGNALGIIGSGS 374
QY 369 GKSSLLRGIVVPCAAVILRDCVIDIKOMPEKIGRIVGLPODIEFSTGVQONLARE 428
Db 375 GKSTILARLLTGAAMPBAKVRLDGASPDQMBRETLRIGLIPGVLEFPDSTIGENISIR 434
QY 429 TEFESQE-VIEAATLAGVHEMIQSLPMGYDVAIEGASLSGGOROLATARAVERFPAL 487
Db 435 DDNPPEALIIAARAAGHEILIREKGYDSDIEBAGSTLSAAGORRIGLARALYGPFI 494
QY 488 LVLDPEPNASLDQVGEVALIMEAMKRLKAARKTVIFATHRKVNLLAODAYIMVINOVISDFG 547
Db 495 VLDPEPNALIDAEGLAAVYVKAITSYKARGIAVVVAHPRPASIAGVADYIILMEDGRMAFG 554
QY 548 ERD 550
Db 555 PRD 557

```

RESULT 10  
AF2972  
Rhizobium secretion [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AF2972  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woos-  
erage, G.; Gillett, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McCl-  
rard, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam-  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AF2972  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-615 <KUR>  
A:Cross-references: GB:AE008689; PIRN:AA14196.1; PID:g17741774; GSPDB:GN00187  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: pTSD  
A:Map position: linear chromosome

Query Match Similarity 36.7%; Score 1029; DB 2; Length 615;  
Best Local Similarity 39.2%; Pred. No. 2, 2e-65;  
Matches 213; Conservative 118; Mismatches 196; Indels 16; Gaps 5;

```
OY      22 AVTTAMVFSEF-----INITALVSPLYMLQVDRYLTSRNVSTLIVTVICVFLV 73  
       ||::: : ::||| |::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     17 AVVSPLKSTFLGVALVSGVNILLATLTGLFMQLGYDRLVAGSLPTLGLALLGLGYGF 76  
  
OY      74 YGLLELRQVLVRGLKFP----GVARDPIFKSVLDSTLSRKGIIGCGAIFPMDOVREM 129  
       ||::: : ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     77 QCLDLIRARVLRIGEDDFMRYSGRVHDADVRLPLVNRM--PGGDLPRLRLDNVRGFL 134  
  
OY     130 TG-GLIAFDAPRTPEFVIYSWMLHPFFGLAIITACIIFIIGLVANDNATKNPDIOMATYA 188  
       |:| | | | | |::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     135 SGTGPFAFFDLPMPPLTLCFLFHVWIGVTALAGAVLLVSLTLITNFISQPIPDNTWE 194  
  
OY     189 STIAONDAGSTLRNAEVMKMGMWGLQARWRARRDEQVANOAAASDAGAVMSGIKYER 248  
       ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     195 NMAARNQLASRRNAEYVQAMGLGRIGKRMQSSNEAYLANRKGADVAAGGLNLAKSLR 254  
  
OY     249 NIYQTIILGGAVLAIDGKISAGAMTAGSILVGRALAPIEGAVGMKNTYGARGMDRLQ 308  
       ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     255 VVLQSAITLVGAWLVRQESAGGYMIASSIMGRALAVDLAISMKSFVARQSVAKLAR 314  
  
OY     309 TMLREKSADDHNPRLPPRCVLSAEAAISIIPGAOPMPROASFRLIDGAALVGPSRA 368  
       ||::: : ::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     315 ELFAQPANANAAMALKRKEELREYNVTIYPGERKKPYTGIGFVYTACNALGITPPSS 374  
  
OY     369 GKSSLRGIVGWPCAGVIRLDGTDIKOWDPEKIGHRVGYLPODIELESFGTAONIAEF 428  
       ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     375 GKSTARILITGAMMPAGKVRLDGASFQDMORETIGRHIGYLPQCVELDPGTGENISR 434  
  
OY     429 TEFEOSD-VTEATITAGVHEMIOSLPMGIDPTALIEGGASLSGGOQRQLALAVFEMPAL 487  
       :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     435 DNNPDEAIIAARAAGAHELILFEKGYDSDIGAGSTLSAGQRIGLARLYGDPL 494  
  
OY     488 LVYDENASILDQGEVNLMEAMKRLKARTVIFETHVNNLQAODYIVINVGYSIDEG 547  
       :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
Db     495 VLIDPRANLIDAGEAAYVKAITSKARGIAVVAAHRPSAIGAVDYILMEDGRKARF 554  
  
OY     548 ERD 550  
       ||  
Db     555 PRD 557
```

RESULT 11  
S12525  
metalloproteinas ABC-type transport system ATP-binding protein prtd [validated] - Erwinia chrysanthemi  
N:Alternate names: metalloprotease export system ABC-type transport component  
N:Contains: adenosinetriphosphatase (EC 3.6.1.3)  
C:Species: Erwinia chrysanthemi  
C>Date: 21-Nov-1993 #sequence\_revision 21-Jul-1995 #text\_change 02-Feb-2001  
C:Accession: S12525; T48666  
R:Letofer, S.; Delepeleire, P.; Wandersman, C.  
EMBO J. 9, 1375-1382, 1990  
A:Title: Protease secretion by Erwinia chrysanthemi: the specific secretion functions are  
A:Reference number: S12524; M01D:90228333  
A:Accession: S12525  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-575 <LEFT>  
A:Cross-references: GB:X53253; NID:g41546; PIDN:CNA37342.1; PID:g581096  
R:Delepeleire, P.; Wandersman, C.  
submitted to the EMBL Data Library, April 1991  
A:Description: C-terminal secretion signal in protease secreted by gram negative bacteria  
A:Reference number: Z24521  
A:Accession: T48666  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-575 <DEL>  
A:Cross-references: EMBL:M60395; PIDN:AAA63634.1  
A:Experimental source: EMBL strain B374  
C:Genetics:

[illegible]





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:42:20 ; Search time 17.51 Seconds  
(without alignments)  
1227.262 Million cell updates/sec

Title: US-09-913-414-4  
Perfect score: 2807  
Sequence: 1 MFRRSGAKPTFDQAVLVAR.....MNINGVISDPGENDRCWPS 555

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt-40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1231	43.9	593	1 APRD_PSEAE	003024 pseudomonas
2	1026.5	36.6	575	1 PRTD_ERWCH	P23596 erwinia chr
3	584.5	20.8	707	1 HLYB_ACTAC	P23702 actinobacill
4	579.5	20.6	708	1 HLYB_PASHA	P16532 pasteurella
5	576.5	20.5	708	1 HLYB_PASAP	P55122 pasteurella
6	569.5	20.3	707	1 RTIB_ACTPL	P26760 actinobacill
7	562.5	20.0	707	1 HLYB_PROVU	P11599 proteus vul
8	561.5	20.0	707	1 HLYB_ECOLI	P10089 escherichia
9	561.5	20.0	707	1 HLYB_ECOLI	P08716 escherichia
10	552.5	19.7	711	1 RT3B_ACTPL	004473 actinobacill
11	533	19.0	712	1 CVAB_BORPE	P18770 bordetella
12	453	16.1	842	1 ABC6_HUMAN	Q9N588 homo sapien
13	445.5	15.9	586	1 Y4G6_RHISN	P53469 rhizobium s
14	440.5	15.7	604	1 YFIC_BACSU	P54718 bacillus su
15	440	15.7	582	1 MSBA_ECOLI	P27299 escherichia
16	419.5	14.9	587	1 MSBA_HAEIN	P44407 haemophilus
17	415	14.8	590	1 MDIA_ECOLI	P77265 escherichia
18	410	14.6	722	1 MESD_LEUME	Q10418 leucostoc
19	408	14.5	573	1 YFIB_BACSU	P54718 bacillus su
20	404	14.4	735	1 ABC8_HUMAN	Q9N412 homo sapien
21	401.5	14.3	694	1 ABC7_MOUSE	Q61102 mus musculu
22	401.5	14.3	752	1 ABC7_HUMAN	Q75027 homo sapien
23	399	14.2	690	1 ATMI_YEAST	P40416 saccharomyc
24	394.5	14.1	631	1 YC72_MYCTU	Q11447 mycobacteri
25	394.5	14.1	1321	1 MDR1_CAEBL	P34712 caenorhabdi
26	392.5	14.0	715	1 LCCL_LACLA	Q9CJ18 lactococcus
27	391.5	13.9	598	1 Y288_THEMEA	Q9WY44 thermocoga
28	390.5	13.9	698	1 CVAB_ECOLI	P25220 escherichia
29	388.5	13.8	859	1 YDA8_MYCTU	Q11018 mycobacteri
30	386	13.8	583	1 EXP8_STRPN	P35598 streptococc
31	386	13.8	717	1 COMA_STRPN	Q03727 streptococc
32	380.5	13.6	1279	1 MDR3_HUMAN	P21339 homo sapien
33	380.5	13.6	1905	1 TAGB_DICDI	P54683 dictyostell

34	379.5	13.5	1254	1 MDR3_CAEBL	P34713 caenorhabdi
35	379.5	13.5	1280	1 MDR1_HUMAN	P08183 homo sapien
36	378.5	13.5	715	1 LCNC_LACLA	000564 lactococcus
37	377	13.4	693	1 YETI_SCHPO	Q14286 schizosacch
38	376.5	13.4	1276	1 MDR2_CRIGR	P21449 cricetus
39	375.5	13.4	575	1 YWA_BACSU	P45861 bacillus su
40	374.5	13.3	1281	1 MDR3_CRIGR	P23174 cricetus
41	369.5	13.2	1743	1 TAGC_DICDI	Q23688 dictyostell
42	367.5	13.1	1276	1 MDR1_CRIGR	P21448 cricetus
43	366.5	13.1	1276	1 MDR2_MOUSE	P21440 mus musculu
44	366.5	13.1	1278	1 MDR2_RAT	Q08201 rattus norv
45	365.5	13.0	584	1 LMRA_LACLA	Q9ch18 lactococcus

ALIGNMENTS

RESULT ID	APRD_PSEAE	STANDARD	PRT	593 AA.
AC	Q03024;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Alkaline protease secretion ATP-binding protein aprd.			
GN	APRD OR PA1246.			
OS	Pseudomonas aeruginosa.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OX	Pseudomonas.			
NCBI_TaxID=287;				
[1]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 15692 / PA01;			
RX	MEDLINE-20437337; PubMed-10984043;			
RA	Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,			
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Garber R.L., Goltzy S., Tolentino E., Westbrock-Wadman S., Yuan Y.,			
RA	Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PA01, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
CC	- FUNCTION: INVOLVED IN THE SECRETION OF ALKALINE PROTEASE.			
CC	- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; X64558; CAA45855.1; -			
DR	EMBL; AE004554; AAG04635.1; -			
DR	PIR; S26696; S26696.			
DR	HSSP; P13569; INBD.			
DR	InterPro: IPR003593; AAA.			
DR	InterPro: IPR001140; ABC_transporter_tmem.			
DR	InterPro: IPR003439; ABC_transporter.			
DR	InterPro: IPR001687; ATP_GTP_A.			
DR	Pfam; PF00664; ABC_membrane; 1.			



DR Pfam: PF00005; ABC\_tran; 1.  
DR SMART: SM00382; AAA; 1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
KW Transmembrane; Transport; ATP-binding; Complete proteome.  
FT NP\_BIND 366 373 ATP (BY SIMILARITY).  
SQ SEQUENCE 593 AA; 63670 MW; CA3AB17BCC27318 CRC64;

Query Match 43.9%; Score 1231; DB 1; Length 593;  
Best Local Similarity 46.5%; Pred. No. 1,le-77;  
Matches 251; Conservative 109; Mismatches 178; Indels 2; Gaps 2;

QY 14 QAVLVAPVITAMVFEFFINILALVSPLYMLQYDRLTSRNVSTLIVLVICVFLFLV 73  
DB 14 QALASGALRSVAAGFSGVILMLVPSLYMLQYDRLTSRNVSTLIVLVICVFLFLV 73  
QY 74 YGLLEALRTQVLRGGLKFCQVARDPIFKSVLDSTLSRKG-IGQARFDMQVDFMTG- 131  
DB 74 MGALAEALRSFVLVRSERFDQHLGRIVAAFEENLRAGGGEASQALHDLTTLKQFTGQ 133  
QY 132 GLIACDAPVTPVIVISWMLHPFEGILATITIFGLAVMNDATKNPITOMATMASIA 191  
DB 134 ALFAPFAPVPPVVLIVLFLFDPMLGLISLVGALALMALAFNERATRAPLAKAGLSIK 193  
QY 192 AQNDAGSTLRNAEYWKAMGWLQARARARDEQVAMQAAASDAGAVMSGIVFRNTV 251  
DB 194 SGQLASNNLRNAEVIEMNGMGLSKRWERLHQAFLDQOSLASRAARINLSKYLRIAL 253  
QY 252 QTLILGGATYLAIDKTSAGAMINGSTLVGRALPIEGAVGQMKNYIGARAMDRLQTM 311  
DB 254 QSLVLGAGWALVAGRTTPGMINGSTILMGALPDIQILGVWQWGAARAYRRLSGL 313  
QY 312 REEKSDHMPLEPPRGVLSAEASILPPGAQOPTMROASFRIDAGAAVALVGSAAKS 371  
DB 314 DEFARERRMELPPRGHLLLESIDAPRSEARTLGLTALTRAGSVVGYIGSGSKS 373  
QY 372 SLRGIWGPCAAGVIRLDGYDIKQMDPEKLRHNVGLIPODIELFSGTVQNTARTFE 431  
DB 374 SLAVVLGIMPTLHGSVLDGAEIRQYERETLGRIGVLPDIELFAGTVAMENIARFE 433  
QY 432 ESQVIEATATAGVHEMOSLPMGYDPAIGSGASLSGGQORLALRAVRRMALLVLD 491  
DB 434 QADVVEARLARGVHELVLRIPQGDIVLVGAGLSGGQORLALALGAPLVVLD 493  
QY 492 EPNASLDQVSEVALMEAKRLKAARVYIFATHRVNLIAQADYIMVINOQYISDFEGR 551  
DB 494 EPNNSLDQSGEALLAIAQAKARGCYLLITHRAGVLCGADRLLALNAQGLHLGERDQ 553

RESULT 2  
PRD\_ERMCH  
ID PRD\_ERMCH STANDARD; PRT; 575 AA.  
AC P23596;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Proteases secretion Atp-binding protein prtd.  
GN PRD.  
OS Erwinia chrysanthemi.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Pectobacterium.  
OX NCBI\_TaxID=556;  
RN [1]  
RP SEQUENCE FROM N.A. PubMed=2184029;  
RA Letoffe S., Delepeatre P., Wandersman C.;  
RT "Protease secretion by Erwinia chrysanthemi: the specific secretion  
RT functions are analogous to those of Escherichia coli  
RT alpha-haemolysin."  
RL EMBO J. 9:1375-1382(1990).  
CC -!- FUNCTION: INVOLVED IN THE SECRETION OF PROTEASES A, B, C AND G.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: M60395; AAA63634.1; -  
DR EMBL: X53253; CA37342.1; -  
DR PIR: S12525; S12525.  
DR HSP: P13569; INBD.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmem.  
DR InterPro: IPR003439; ABC\_transportr.  
DR InterPro: IPR001687; ATP-GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 1.  
DR SMART: SM00382; AAA; 1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
KW Transport; ATP-binding; Transmembrane; Inner membrane.  
FT TRANSMEM 22 42 POTENTIAL.  
FT TRANSMEM 56 76 POTENTIAL.  
FT TRANSMEM 133 153 POTENTIAL.  
FT TRANSMEM 154 174 POTENTIAL.  
FT TRANSMEM 255 275 POTENTIAL.  
FT NP\_BIND 364 371 ATP (BY SIMILARITY).  
SQ SEQUENCE 575 AA; 61617 MW; 1BCBC4E74B1B92D6 CRC64;

Query Match 36.6%; Score 1026.5; DB 1; Length 575;  
Best Local Similarity 43.5%; Pred. No. 1,4e-63;  
Matches 225; Conservative 94; Mismatches 193; Indels 5; Gaps 4;

QY 28 VFSEFFINILALVSPLYMLQYDRLTSRNVSTLIVLVICVFLFLVGLLEALRTQVLR 87  
DB 27 IFSAVINVLMLAPSVMLQYDRLTSRNVSTLIVLVICVFLFLVGLLEALRTQVLR 86  
QY 88 GGLFEDVARDPIFKSVLDSTL-SRKGIGQAFRDMQVDFMTG-GLIACDAPVTPV 145  
DB 87 LGTRIDALNDQVNAAFARLLEAGDGRAGLALTDLLKROFTIGNALFAFDVPPWPLF 146  
QY 146 VIVSWMLHPFEGILATITIFGLAVMNDATKNPITOMATMASIAQNDAGSTLRNAEV 205  
DB 147 LVLFLHPMLGMLALGTATVPGGVGLAEPASDQSTAGSNOGQATHLADQLRADV 206  
QY 206 MKAMGNGGLQARARARDEQVAMQAAASDAGAVMSGIVFRNTVOTLILGGATYLAID 265  
DB 207 IEAMGMLGNLRRWLARHYRFLSLQNLASERRAAVAGASKYSRTALQSLMGLGALLAID 266  
QY 266 GKISGAMINGSTLVGRALPIEGAVGQMKNYIGARAMDRLQTMREEKSADHMPLE 325  
DB 267 GKITPGMIMIASILVGRVSLPDIQILGVWQWGAARAYRRLSGL 326  
QY 326 PRGVLSAEASILPPGAQOPT-MROASFRIDAGAAVALVGSAAKSSLRGIVGPCA 384  
DB 327 PEGHLSDQVSLFR-AQGNTRLQNIHPSLQAGTLVILGSGSGKSLARLLVGAQSP 384  
QY 385 AGVIRLDGYDIKQMDPEKLRHNVGLIPODIELFSGTVQNTARTFESEQVIEATL 444  
DB 385 QGKVRLDGADLNDQVKNFTPTIGYLDVDFGSLAENIARFGADPEKVVAAKLAG 444  
QY 445 VHEMIOSLPMGYDPAIGSGASLSGGQORLALRAVRRMALLVLDPEPNASLDQVEVA 504  
DB 445 VHEILSLPNCYDTELGGGGGSLGGQORGLRAMGDPCLILDEPNASLDQVEVA 504  
QY 505 LMEAMKRLKAARVYIFATHRVNLIAQADYIMVINOQ 541  
DB 505 LMQAIVALQKRGATVVLITHRALPITLLAOKLILHEG 541

RESULT 3



```
HLVB_ACTAC
ID HLVB_ACTAC STANDARD: PRT: 707 AA.
AC P23702.
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukotoxin secretion ATP-binding protein.
GN LKTB OR LTB.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OC actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_TaxID=714.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90384842; PubMed=2402457;
RA Guttmiller J.M., Kolodrubetz D., Cagle M.P., Kraig E.;
RT "Sequence of the lktB gene from Actinobacillus
actinomycetemcomitans."
RL Nucleic Acids Res. 18:5291-5291(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92072446; PubMed=1961107;
RA Lally E.T., Golub E.E., Kleba I.R., Taichman N.S., Decker S.,
Berthold P., Gibson C.W., Demuth D.R., Rosenbloom J.;
RT "Structure and function of the B and D genes of the Actinobacillus
actinomycetemcomitans leukotoxin complex."
RL Microb. Pathog. 11:111-121(1991).
CC -1- FUNCTION: INVOLVED IN THE EXPORT OF LEUKOTOXIN (HEMOLYSIN).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLVB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53955; CA37906.1; -.
DR PIR; S12601; S12601.
DR PIR; A61378; A61378.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmem.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam; PF00064; ABC_membrane; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hemolysis; Transport; ATP-binding; Transmembrane.
FT NP_BIND 502 509 ATP (BY SIMILARITY).
SQ SEQUENCE 707 AA; 79578 MW; CA564EACF46DF4FB CRC64;
-----
Query Match 20.8%; Score 584.5; DB 1; Length 707;
Best Local Similarity 30.0%; Pred. NO.5.4e-33;
Matches 169; Conservative 106; Mismatches 247; Indels 41; Gaps 12;
```

```
Db 321 DKFARNADNOSFLVESYAIN-----TIKAMAIPOQWTNIM----DKOLASYAV 366
Qy 234 SDAGAV-----MSGIKVRNRYQVLTLLGGAVLADGKISAGMAGSILYGRALAPLEG 289
Db 367 SEKYVTLATIGQOOGIQLQKAVMINTLWGAHLVIGSLSTIGLAFNMLAGQIISPVR 426
Qy 290 AVGWKNYIGARGAMDRIQTMLR--EKSADDDHPLPEPRGVLSAFAASTLPGAQOPT 346
Db 427 LAQIMQDFQVIGISTRLGDVNLSTENNTAS--VSLPEIGSEISFRNKKFYKDSFMT 484
Qy 347 MROASFRIDAGAAVALVPSAAGKSSLLRGIVGWPACAGVIRLDGYDIKQMPKELRN 406
Db 485 LNNIMLDISQGEVIGIVGRSSGKSTLRKLRQREYTPPGGVLDIGDHLADPWWLRQ 544
Qy 407 VGYLPDIELEFSGVYAONIARTE-FESQEVTEATLGVHEMIOSLPMGVDTAEGGA 465
Db 545 VGVVLDQDVNLNRSIRENIALTNPCMPMEKVIYAAAKLAGAHDFISELEGEYNTVVEOGA 604
Qy 466 SLSSGQORLALRAVFRMPALVLDPEPNASLDQGEVALMEAMRKLRKAARTVIFATHK 525
Db 605 GLSGQORRIARALVNNPRLILPDEATSDLDYSENIIMHNMHT-CQNTVLIIAHR 663
Qy 526 VNLAQADYIMVINGVISDFGE 548
Db 664 LSTVKNADRIITVMDKGETIEQK 686

RESULT 4
HLVB_PASHA
ID HLVB_PASHA STANDARD: PRT: 708 AA.
AC P16532.
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukotoxin secretion ATP-binding protein.
GN LKTB.
OS Pasteurella haemolytica.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Mannheimella.
OX NCBI_TaxID=75985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE A1.
RX MEDLINE=87306837; PubMed=3040588;
RA Lo R.Y.C., Strathdee C.A., Shewen P.E.;
RT "Nucleotide sequence of the leukotoxin genes of Pasteurella
haemolytica A1."
RL Infect. Immun. 55:1987-1996(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE A1.
RX MEDLINE=89123172; PubMed=2914876;
RA Strathdee C.A., Lo R.Y.C.;
RT "Cloning, nucleotide sequence, and characterization of genes encoding
the secretion function of the Pasteurella haemolytica leukotoxin
determinant."
RL J. Bacteriol. 171:916-928(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE A1 / PHU101.
RX MEDLINE=89210283; PubMed=2707120;
RA Highlander S.K., Chidambaram M., Engler M.J., Weinstein G.M.;
RT "DNA sequence of the Pasteurella haemolytica leukotoxin gene
cluster."
RL DNA 8:15-28(1989).
RN [4]
RP SEQUENCE OF 1-48 FROM N.A.
RX MEDLINE=90236888; PubMed=2185213;
RA Highlander S.K., Engler M.J., Weinstein G.M.;
RT "Secretion and expression of the Pasteurella haemolytica leukotoxin."
RL J. Bacteriol. 172:2343-2350(1990).
CC -1- FUNCTION: INVOLVED IN THE EXPORT OF LEUKOTOXIN.
```

```
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR EMBL; M20730; AAA25530.1; -;  
DR EMBL; M24197; AAA25544.1; -;  
DR PIR; A32051; A32051.  
DR PIR; S29517; S29517.  
DR InterPro; IPR003593; AAA.  
DR InterPro; IPR001140; ABC\_transporter\_tmem.  
DR InterPro; IPR003439; ABC\_transportr.  
DR InterPro; IPR001687; ATP\_GTP\_A.  
DR Pfam; PF00664; ABC\_membrane; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
KW Hemolysis; Transport; ATP-binding; Transmembrane.  
FT TRANSMEM 216 236 POTENTIAL.  
FT TRANSMEM 249 269 POTENTIAL.  
FT TRANSMEM 327 347 POTENTIAL.  
FT TRANSMEM 353 373 POTENTIAL.  
FT TRANSMEM 446 466 POTENTIAL.  
FT TRANSMEM 468 488 POTENTIAL.  
FT NP\_BIND 503 510 ATP (BY SIMILARITY).  
FT NP\_BIND 512 512 Q -> T (IN REF. 3).  
FT CONFLICT 340 340 S -> A (IN REF. 3).  
FT CONFLICT 365 367 SS -> FI (IN REF. 3).  
FT CONFLICT 597 597 N -> T (IN REF. 3).  
SQ SEQUENCE 708 AA; 79712 MW; BE2ABAF12380A4E CMC64;

Query Match 20.6%; Score 579.5; DB 1; Length 708;  
Best Local Similarity 29.6%; Pred. No. 1.2e-32;  
Matches 166; Conservative 116; Mismatches 243; Indels 35; Gaps 14;

QY 10 TTFDOAVLPAVITAMVSEFFINITALVSPYMLQYVDRLTSRNVSTLVLYVYICVF 69  
DB 142 TWEIFAVIKYRKIFETLIVSIFLQIFALITPLEFOVYMDKVLVHGFSTLNIITVALAI 201  
QY 70 LFLVYGLLEALTOYLVNRGLKFGCVARDPIKSVLDSTLS-----RKIGGAQAFDMQ 124  
DB 202 VLIETIVISGLRTYVSHSTSRIDVELGAKLFRHLISLPISFENRRVGDIVARRELDQ 261  
QY 125 VEEFMTG-GLIAFCDAPTVPEVIVSWMLHP--FFGILAIACIIIFGLAVNNDATKNP 181  
DB 262 TINFLEGLQALTSVLDLTFEIFFAWMYSPKLTIVLIGSLCYILMSIFI-----SP 314  
QY 182 I---QMATMASTAAQND--GSTLRNAEYKAMGCGLOARRRARDEQVMAQAAASA 236  
DB 315 IIRRLDEKFAASADNOAPLVESVTSINNIKAMAAVAPQMTDW---DKOLASVSSSPR 370  
QY 237 GGAIV-----MSGIKVFNIVYQTLILGGAYLAIDKISAGAMAGSTLVGRALPIEGANG 292  
DB 371 VYVLATIGGQGVOLQKTYKVINLWLGALVLSGDLSIQLLAFNLSQVLAIVRLAQ 430  
QY 293 QMKNYIGARGAMDRLQTMLE- EKASADHMLPEPRGVLSAEAAST-LPPGAQQPT-MRQ 349  
DB 431 LMDFQGVGISTVRLGDLVNSPTEQYQKLSLPEIKGDISFKNIRRRYPDA--PILNN 488  
QY 350 ASFRIDAGAAVALVGPASAGKSSLLRGIVGWPCCAAGVIRLDYDIKQMDPEKLGHVGY 409  
DB 489 VMLIEIRGEGVIGIVSGSGSKSTLKILORFYIPENGQVLIIDHDALADPMMLRQIGV 548  
QY 410 LQDIEFLFGVYQANTARTE- FESEVIEAATLAVHEMIGSLPBGYDTAIGEGASIS 468  
DB 549 VLDNVLLNRSIRENTALSDPGMPMERVITYAAKLAGADHITSELREGVITIVEGGAGIS 608

QY 469 GGGORIALARAFRPMALLVDEPNASLDQGVVALMEAMKRLAKRTVIFATHKVNL 528  
DB 609 GGGORIALARAFRPMALLVDEPNASLDQGVVALMEAMKRLAKRTVIFATHKVNL 667

QY 529 LAQDYIMVINOYISDFGE 548  
DB 668 VKNADRIIVMEKGEIVGQK 687

RESULT 5  
HLYB\_PASSP STANDARD; PRT; 708 AA.  
ID HLYB\_PASSP  
AC P55122;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Leukotoxin secretion ATP-binding protein.  
GN LKTB.  
OS Pasteurella haemolytica-like sp. (strain 5943B).  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=28165;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93239320; PubMed=8478098;  
RA Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;  
RT "Molecular characterization of a leukotoxin gene from a Pasteurella  
RT haemolytica-like organism, encoding a new member of the RTX toxin  
RT family";  
RL Infect. Immun. 61:2089-2095(1993).  
CC -1- FUNCTION: INVOLVED IN THE EXPORT OF LEUKOTOXIN.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L12148; AAA16445.1; -;  
CC InterPro; IPR003593; AAA.  
CC InterPro; IPR001140; ABC\_transporter\_tmem.  
CC InterPro; IPR003439; ABC\_transportr.  
CC InterPro; IPR001687; ATP\_GTP\_A.  
CC Pfam; PF00664; ABC\_membrane; 1.  
CC Pfam; PF00005; ABC\_tran; 1.  
CC SMART; SM00382; AAA; 1.  
CC PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
KW Hemolysis; Transport; ATP-binding; Transmembrane.  
FT TRANSMEM 159 176 POTENTIAL.  
FT TRANSMEM 189 212 POTENTIAL.  
FT TRANSMEM 270 288 POTENTIAL.  
FT TRANSMEM 295 316 POTENTIAL.  
FT TRANSMEM 388 404 POTENTIAL.  
FT TRANSMEM 411 428 POTENTIAL.  
FT NP\_BIND 503 510 ATP (BY SIMILARITY).  
SQ SEQUENCE 708 AA; 79772 MW; 7326D21411B5090A CMC64;

Query Match 20.5%; Score 576.5; DB 1; Length 708;  
Best Local Similarity 29.2%; Pred. No. 1.9e-32;  
Matches 165; Conservative 116; Mismatches 243; Indels 41; Gaps 14;

QY 10 TTFDOAVLPAVITAMVSEFFINITALVSPYMLQYVDRLTSRNVSTLVLYVYICVF 69  
DB 142 TWEIFAVIKYRKIFETLIVSIFLQIFALITPLEFOVYMDKVLVHGFSTLNIITVALAI 201  
QY 70 LFLVYGLLEALRTQYLVNRGLKFGCVARDPIKSVLDSTLS-----RKIGGAQAFDMQ 124

Db 202 VIREIYLSGLRTYIFAHSTSRIDVELGARLFRLHLLPISYFENRRVGDVAVAREIDQ 261  
QY 125 VREMTG-GLIAFCDAWPTPEVIVSMMLHP--FEGILAIACIIINGLAN-----MN 174  
Db 262 IRNFLTGQALTSVLDLMSFFFAVMWYSPKLTVLIGSLPCYILMSIFSPILRRRLD 321  
QY 175 DNATKRNIGMATMA-SIAONDAGSTLRNAEVMKMGWGLQARWRARDEQVAMQAAA 233  
Db 322 DKFRAGNDQNAFLVESVTAIN-----MIKMAVSPQMTWTW----DKOLASVSS 367  
QY 234 SDAGAGV----MSGIKVFRNIYOTLLIGGAYLAIDKISAGAMIAGSILVGRALAPLEG 289  
Db 368 SFRVTYATIGQGVQVLOKTVWYINLMGLAHLYISGDLISGOLIAFNMLSGQVYAPYIR 427  
QY 290 AVGOMKRYIARGAMDRLOQMLRE-EKSADDMPLPEPRGVLSAEAAST-LPPAQOQPT- 346  
Db 428 LAOLMDFOTGVGISVTLGDVLSPTBOYOGKLSLPEIQGDIARKNIRFRYKPPA--PTI 485  
QY 347 MRQASFRIDAGAAVALVPSAAGKSSILRGIVGPCAAGVIRLDGYDKOMDEKLGRIH 406  
Db 486 LNNVNLTKKGEVIGIVRGSGSGKSTLTLLKLRFTYIPENGVLIDGHDALADNMLRRQ 545  
QY 407 VGVLPDIELEPSTVAONIAFTEESOE-VIEATLAGVHEMTQSLPMGYDAIGEGGA 465  
Db 546 IGVVLQDNVLNRSIRENIALSEPGMSMERIYNAKLAGAHDFISDVEGNYITIGEGGA 605  
QY 466 SLGGGQORLALAVFRMALLVLDENPNSLDVGEVALMEAKRKAKRYTIFATHK 525  
Db 606 GLSSGGQORLALAVFRMALLVLDENPNSLDVGEVALMEAKRKAKRYTIFATHK 664  
QY 526 VNLQAQDYIMVINOQVYISDFGERD 550  
Db 665 LSTYKNADRIIVMEKGEIVGEGKHN 689

RESULT 6  
RTLB\_ACRPL STANDARD; PRT; 707 AA.  
ID RTLB\_ACRPL  
AC P26760;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE RTX-I toxin determinant B (Toxin RTX-I secretion ATP-binding protein)  
DE (APX-IB) (CLY-IB) (CYCLOSLIN IB) (CLY-IB).  
GN APXIB OR CLYIB OR HLYIB OR APPB.  
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;  
OC Actinobacillus.  
OX NCBI\_TaxID=715;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SEROTYPE 5;  
RX MEDLINE-91317735; PubMed-1860823;  
RA Chang Y.-F., Young R., Struck D.K.;  
RT "The Actinobacillus pleuropneumoniae hemolysin determinant: unlinked  
RT appca and appbd loci flanked by pseudogenes.";  
RL J. Bacteriol. 173:5151-5158(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ISOLATE CVI 13361 / SEROTYPE 9;  
RX MEDLINE-92040145; PubMed-1937809;  
RA Smith M.A., Briatore J., Jansen R., Smith H.E., Kamp E.M.,  
RA Gielkens A.L.J.;  
RT "Glycolysins of Actinobacillus pleuropneumoniae serotype 9.";  
RL Infect. Immun. 59:4497-4504(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S 4074 / SEROTYPE 1;  
RX MEDLINE-94237497; PubMed-8181764;  
RA Frey J., Haldemann A., Nicolet J., Boffini A., Prentki P.;  
RT "Sequence analysis and transcription of the aptx operon (hemolysin I)  
RT from Actinobacillus pleuropneumoniae.";

RL Gene 142:97-102(1994).  
CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-I AS WELL AS  
CC THAT OF RTX-II.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 9.  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M65808; AAB00966.1; -;  
DR EMBL; X61112; CAA43425.1; -;  
DR EMBL; X68595; CAA48587.1; -;  
DR PIR; A40366; A40366.  
DR PIR; S18855; S18855.  
DR InterPro; IPR003593; AAA.  
DR InterPro; IPR001140; ABC\_Transporter\_tmem.  
DR InterPro; IPR003439; ABC\_transportr.  
DR InterPro; IPR001687; ATP\_GTP\_A.  
DR Pfam; PF00664; ABC\_membrane.1.  
DR Pfam; PF00005; ABC\_tran.1.  
DR SMART; SM00382; AAA.1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER.1.  
KW Hemolysis; Cytolysis; Transport; ATP-binding; Transmembrane.  
FT TRANSMEM 158 179 POTENTIAL.  
FT TRANSMEM 187 204 POTENTIAL.  
FT TRANSMEM 269 289 POTENTIAL.  
FT TRANSMEM 293 311 POTENTIAL.  
FT TRANSMEM 387 403 POTENTIAL.  
FT NP\_BIND 502 509 ATP (BY SIMILARITY).  
FT VARIANT 110 112 EOA -> KQT (IN SEROTYPE 5).  
FT VARIANT 328 328 D -> H (IN SEROTYPE 5).  
FT VARIANT 517 517 R -> V (IN SEROTYPE 5).  
FT VARIANT 556 556 N -> G (IN SEROTYPE 5).  
FT VARIANT 609 611 GQR -> RAT (IN SEROTYPE 1).  
FT VARIANT 611 612 DR -> PN (IN SEROTYPE 5).  
FT VARIANT 671 672 DR -> AS (IN SEROTYPE 5).  
FT VARIANT 692 692 A -> R (IN SEROTYPE 5).  
SQ SEQUENCE 707 AA; 79663 MW; 34A4339C57340DP9 CRC64;

Query Match 20.3%; Score 569.5; DB 1; Length 707;  
Best Local Similarity 28.4%; Pred. No. 5.8e-32;  
Matches 161; Conservative 105; Mismatches 253; Indels 47; Gaps 9;

QY 10 TIFQAVLVNAPAVITAMVFSFINILALVSPMLQVYDVLVSRNVSTILVTVICVF 69  
Db 141 TWIIPAVIKRKIFETLIVSIFLQIFALITPLEFYQVMDVLVHREGSTLVNTVALAI 200  
QY 70 LFLVYGLLEALRTQVIVRGGLKFGVARDPIFKSLVDSTLS-----RKIGIGQAFRWDQ 124  
Db 201 VLEIYVINGLRTYIFAHSTSRIDVELGARLFRLHLLPISYFENRRVGDVAVAREIDQ 260  
QY 125 VREMTG-GLIAFCDAWPTPEVIVSMMLHPFEGILAIACIIIFGLAVMNDNAFKNPIDQ 183  
Db 261 IRNFLTGQALTSVLDLMSFFFAVMWYSPKLTVLIGSLPCYILMSIFSPILRRRLD 320  
QY 184 MATMASTIAONDAGSTLRNAEVMKMGWGLQARWRARDEQVAMQAAA-----SDAGGA 239  
Db 321 EKFRAGNDQNSFLVESVTAINITKALAVTPQMTWTW----DKOLASVSAEFRVTTLATI 376  
QY 240 VMSGIKVFRNIYOTLLIGGAYLAIDKISAGAMIAGSILVGRALAPLEGAVGOMKNVIG 299  
Db 377 GQGVQVQTKQVWYITMLGLAHLYISGDLISGOLIAFNMLSGQVYAPYIRLAOLMDFQOQ 436  
QY 300 ARGAMDRLOQMLRE-EKSADDMPLPEPRGVLSAEAASTLPPAQOQPTMRQASFR----- 353

```

Db 437 VGISTRTAGDVLSNPSTESYGKALPEIKGDI-----TFNRIRRRYKPA 481
Oy 354 -----IDAGANAALVGPSPAAGKSSLRGIVGWPCAAGYIRLDGYIKOMDEPKL 403
Db 482 PYILDVNLNLSIOOGSEVIGDVIAGRSGSKTLTKLQRFYIPENGQVLLDGHDLALADPNWL 541
Oy 404 GRHVGYLPDIELEFGTAVONIAFRTE-FESOEVIETATLGVHEIMIOSLPMGCDTAIGE 462
Db 542 RQGVSVVLIQDVNLNRSIRDNIALADPCGMPEAKIYHAAKGAAHFEISLEDEGVTIYGE 601
Oy 463 GGASISGQROROLALARAVERFMPALLVLDEPNASIDQVGEVALMEAMKRLKAARTVIPA 522
Db 602 OGAGISGQRORIARALAVNNPKILIFDEATSAIDYESEHIMRNHQI-CKGRVTYIII 660
Oy 523 THKVVLLAQADYIMYINGVISDPGE 548
Db 661 AHRISTVKNADRTIYMERQGIIVEGK 686

RESULT 7
HLXB_PROVU STANDARD: PRT; 707 AA.
AC P11599;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hemolysin secretion ATP-binding protein.
GN HLYB.
OS Proteus vulgaris.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Proteus.
OX NCBI_TaxID=585;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=89039746; PubMed=3054490;
RX Koronakis V., Koronakis E., Hughes C.;
RT "Comparison of the haemolysin secretion protein HlyB from Proteus
RT vulgaris and Escherichia coli; site-directed mutagenesis causing
RT implement of export function.";
RL Mol. Gen. Genet. 213:551-555(1988).
CC -!- FUNCTION: INVOLVED IN THE EXPORT OF HEMOLYSIN A.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC
CC EMBL: X12652; CAA31330.1; -.
CC PIR: S05477; LEEBBV.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmem.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane_1.
DR Pfam: PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hemolysis; Transp; ATP-binding; Transmembrane.
FT TRANSMEM 158 179 POTENTIAL.
FT TRANSMEM 187 204 POTENTIAL.
FT TRANSMEM 269 289 POTENTIAL.
FT TRANSMEM 293 311 POTENTIAL.
FT TRANSMEM 360 377 POTENTIAL.
FT TRANSMEM 387 403 POTENTIAL.
FT TRANSMEM 407 425 POTENTIAL.
FT NP_BIND 502 509 ATP (BY SIMILARITY).
SQ SEQUENCE 707 AA; 79940 MW; 6912C5D4EB18204F CRC64;

```

Query Match 20.0%; Score 562.5; DB 1; Length 707;  
Best Local Similarity 29.7%; Pred. No. 1,8e-31;  
Matches 170; Conservative 102; Mismatches 240; Indels 61; Gaps 14;

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M10133; AAA23976.1; -  
 DR PIR; B24433; LEECB.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmem.  
 DR InterPro: IPR003439; ABC\_transportr.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR Pfam; PF00664; ABC\_membrane; 1.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR SMART; SMO0382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 KM Hemolysis; Transport; ATP-binding; Transmembrane.  
 FT TRANSMEM 158 179. POTENTIAL.  
 FT TRANSMEM 187 204. POTENTIAL.  
 FT TRANSMEM 269 289. POTENTIAL.  
 FT TRANSMEM 293 311. POTENTIAL.  
 FT TRANSMEM 360 377. POTENTIAL.  
 FT TRANSMEM 387 403. POTENTIAL.  
 FT TRANSMEM 407 425. POTENTIAL.  
 FT NP\_BIND 502 509. ATP (BY SIMILARITY).  
 SQ SEQUENCE 707 AA; 79463 MW; 21099CB45E59437E CRC64;

Query Match 20.0%; Score 561.5; DB 1; Length 707;  
 Best Local Similarity 29.7%; Pred. No. 2,Je-31;  
 Matches 172; Conservative 101; Mismatches 233; Indels 73; Gaps 14;

10 TIFQAVLVAPAVITAAVFSEFINIILALVQYDVLNRANSTVLIVLCVF 69  
 141 TWFPATIKYKITEVLIVSVFQLFALITLPEFQVMDKVLVHGRSTLNVITVALSV 200  
 70 LELVYGLLEALRTQVLVNGKLPDGVARDPLFKSVLSTLSL-----RKIGGQAFRMDQ 124  
 201 VVVEFILLISGLRTYIFANSTSRIDVELGAKLFRLHLALPIVFESERRRGDVARRELDQ 260  
 125 VREMTG-GLAFCDAPTPVFVYSMLHP--FFGIIATACIIIFGLAV-----MN 174  
 261 IRNLTGALTSVLDLFSFIFFAVMYSPKLTLYIFLSPCAVAMGVFTSPILRRRLD 320  
 175 DNATKN-PIQATNASTAONDAGSTLNAEVMKMGWGQARWRARBDQVAA 233  
 321 DKFSRNADNOSFLVESYTAI-----TIKAAVSPQNTINW---DKQLAGIYA- 365  
 234 SDAGAV-----MSGIKVERNIVQTLILGGAVLADGKISAGAMTAGSILVGRALAP 286  
 366 --AGFKVTVLATIGQGQIQLQKFTVMINLWGLHVLISGDLSTIGQLAFNMLAGQIYAP 423  
 287 IEGAVGOMKNTIGARGANDRLQTMRE-EKASADHMPPEPRGVLSAANAATLPPGAQOP 345  
 424 VIRLAQIMQDQVQYGISVTRLDVLSNPTESYHGKLAPELNGDI----- 468  
 346 TMRQASR-----IDAGAAVALVGPASAGKSLGLGIYVMPCAAGVRL 390  
 469 TFRNIRFRYKADSPVILLDNILNLSIKGVEYIGIVGRSSGSKSTLTKLIRFTYIPENGOVLI 528  
 391 DGYDIKQMDPEKLRHNGYLPDIDLEFSGTVAQNIARFTEES--QEVTEAATLGVHMI 449  
 529 DGHDLALADPMLRQGVVVLQDNVLLNRSIINDISLANPGMSVSKYIYAKKAGAHDFI 588  
 450 QSLPWGYTAIGEGGASISGQOROLALAVFRMPALLVDEPNVASIDQYGEVALMEAM 509  
 589 SELREGYNTIYEGGAGISGQORRIATARALVNNPKILIFDEATSAIDYSEHVIHNM 648  
 510 KRLAKRTYVFATPKVNLAAQADYIMVINGVSDPGE 548  
 649 KKT--CKGRTVITIAHRLSTVKNADRIIVMERKGIYEGK 686

RESULT 9  
 HLXB\_ECOLI  
 ID HLXB\_ECOLI STANDARD: PRT: 707 AA.  
 AC P08716; 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hemolysin secretion ATP-binding protein, plasmid.  
 GN HLXB  
 OS Escherichia coli.  
 OG Plasmid Inc12 PHLY152.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hess J., Weis W., Vogel M., Goebel W.;  
 RT "Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparison with a corresponding chromosomal hemolysin sequence."; FEBS Microbiol. Lett. 34:1-11(1986).  
 RL [2]  
 RN TOPOLOGY.  
 RX MEDLINE=92204133; PubMed=1552901;  
 RA Gentschev I., Goebel W.;  
 RT "Topological and functional studies on HlyB of Escherichia coli."; Mol. Gen. Genet. 232:40-48(1992).  
 RN [3]  
 RP TOPOLOGY.  
 RX MEDLINE=91132653; PubMed=1994034;  
 RA Wang R.C., Seror S.J., Blight M., Pratt J.M., Broome-Smith J.K., Holland I.B.;  
 RT "Analysis of the membrane organization of an Escherichia coli protein translocator, HlyB, a member of a large family of prokaryote and eukaryote surface transport proteins."; J. Mol. Biol. 217:441-454(1991).  
 RL -1- FUNCTION: INVOLVED IN THE EXPORT OF HEMOLYSIN A.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

FT DOMAIN 409 411 PERIPLASMIC (PROBABLE).  
 FT TRANSMEM 412 432 PROBABLE.  
 FT DOMAIN 433 707 CYTOPLASMIC (PROBABLE).  
 FT NP\_BIND 502 509 ATP (BY SIMILARITY).  
 SO SEQUENCE 707 AA; 79672 MW; 412A3EB64A3CFB4 CRC64;

Query Match 20.0%; Score 561.5; DB 1; Length 707;  
 Best Local Similarity 29.7%; Pred. No. 2,1e-31;  
 Matches 112; Conservative 101; Mismatches 233; Indels 73; Gaps 14;

QY 10 TIFQAVLAVPAVITAMVSEFINILALVSPYLMQVYDRVLTSRNVSTLIVTVICVF 69  
 DB 141 TWFPATIKYRRIFETLVSVFLQFLALITPLEFQVYMDKVLVHRGFSTLVNTVIALSV 200  
 QY 70 LFLVGLLEALRTQVLVVRGKLFQVARDPIFKSVLDSTLS-----RKIGGQAFRDMQ 124  
 DB 201 VVVEIILSGLRITVPSHSTRIDVELGAKLFRHLALPIFESRRKGDVAVARELDQ 260  
 QY 125 VREPMTG-GLIAFCDAWPVVFVIVSWMLHP--FFGILAIACIIIFGLAV-----MN 174  
 DB 261 IRLNLTGQALTSVDLLESLFFAVMWYSPKLTIVILFSLPCYAAMSVFTSPILRRLD 320  
 QY 175 DNATKN-PIQMATASTAOADASTLNAEVMKMGGLQARWRARRDQVAMQAA 233  
 DB 321 DKFSRMDNOSFLVESVAIN-----TIKAMAVSPQMTNIV--DKQLAGIYA- 365  
 QY 234 SDAGAV-----MSGKVFERNIVQTLILGGAVALIDGKISAGAMTSGILVGRALAP 286  
 DB 366 --AGKTVIATIGQGGIQLQKTYMIINLMGALHVLISGDLISQGLAFNMLAQIYAP 423  
 QY 287 IEGAVGQKNTYIGARGAMDRLQTMRE-EKSAADHMLPPEPRGVLSAEASTILPPGAOP 345  
 DB 424 VIRLAQIMQDFQOQVIGISVTRGLDVLSPTESYHGKLTLPETNGDI----- 468  
 QY 346 TMRQASFR-----IDAGAVALVYSPAGSKSLNGIYGVMPGACAGVIRL 390  
 DB 469 TFRNTRFRKPDSPVILDNINLSIKQGEVIGIVSGSGKSTLTKLQIRFYIPKGOVLI 528  
 QY 391 DGYDIKQMDPEKLRHNGVYLPQDIELFSGYVAQNIARFTEES--OEVIETATLAVHEMI 449  
 DB 529 DGHDLALADPNMLRQGVVILQDNVLNRSIIDNISLANPGMSVKYVIAAKLAGAHPEI 588  
 QY 450 QSLPMGYTPTAIGEGGASISGQORLALRAVFRPALIVDEPNASLDQGEVALMEM 509  
 DB 589 SELREGYNTIVIEQCGAGISGQORIALALVNNPKILIEDEASALDYESEHIVRM 648  
 QY 510 KRLKAARVTFATKVNILAQADYIMVINGVISDFE 548  
 DB 649 HRI-CKGRVITIIAHLSTVKMADRIVMEKGKIVEQOK 686  
 RESULT 10  
 RT3B\_ACTPL STANDARD: PRT; 711 AA.  
 ID RT3B\_ACTPL  
 AC Q04473;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE RTX-III toxin determinant B (Toxin RTX-III secretion ATP-binding protein) (APX-III) (Cytolysin IIIB) (CLY-IIIB).  
 GN APXIII OR CLYIII OR RTX.  
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Actinobacillus.  
 OX NCBI\_TaxID=715;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SEROTYPE 2;  
 RX MEDLINE=9326392; PubMed=8494611;  
 RA Chang Y.-F., Shi J.-Y., Ma D.-P., Shin S.-J., Lein D.H.;  
 RT "Molecular analysis of the Actinobacillus pleuropneumoniae RTX toxin-III gene cluster."

RL DNA Cell Biol. 12:351-362(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=405 / SEROTYPE 8;  
 RX MEDLINE=95012630; PubMed=7927703;  
 RA Jansen R., Briare J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J., Smits M.A.;  
 RT "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX) operons: characterization of the ApXIII operons."  
 RT Infect. Immun. 62:4411-4418(1994).  
 RN [3]  
 RP SEQUENCE OF 1-39 FROM N.A.  
 RC STRAIN=405 / SEROTYPE 8;  
 RX MEDLINE=93162836; PubMed=8432615;  
 RA Jansen R., Briare J., Kamp E.M., Gielkens A.L.J., Smits M.A.;  
 RT "Cloning and characterization of the Actinobacillus pleuropneumoniae-RTX-toxin III (ApXIII) gene."  
 RL Infect. Immun. 61:947-954(1993).  
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-III.  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 2.  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLVB SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L12145; AAA21925.1; -;  
 DR EMBL: X80055; CAA56359.1; -;  
 DR EMBL: X68815; CAA48712.1; ALT-INIT.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC transporter\_tmem.  
 DR InterPro: IPR003439; ABC transporter.  
 DR InterPro: IPR001687; ATP-GTP\_A.  
 DR Pfam: PF00664; ABC\_membrane; 1.  
 DR Pfam: PF00005; ABC\_tran; 1.  
 DR SMART: SM00382; AAA; 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 KW Cytolysin; Transport; ATP-binding; Transmembrane.  
 FT TRANSMEM 162 178  
 FT TRANSMEM 197 213  
 FT TRANSMEM 274 290  
 FT TRANSMEM 299 315  
 FT TRANSMEM 391 407  
 FT TRANSMEM 506 513  
 FT NP\_BIND 61 61  
 FT VARIANT 91 91 M -> A (IN SEROTYPE 8).  
 FT VARIANT 91 91 M -> T (IN SEROTYPE 8).  
 FT VARIANT 125 126 GM -> DV (IN SEROTYPE 8).  
 FT VARIANT 213 213 S -> N (IN SEROTYPE 8).  
 FT VARIANT 416 416 T -> A (IN SEROTYPE 8).  
 FT VARIANT 469 469 F -> K (IN SEROTYPE 8).  
 FT VARIANT 606 606 L -> Q (IN SEROTYPE 8).  
 FT VARIANT 638 638 S -> R (IN SEROTYPE 8).  
 SO SEQUENCE 711 AA; 80405 MW; 5B8E46B89D12D92 CRC64;  
 Query Match 19.7%; Score 552.5; DB 1; Length 711;  
 Best Local Similarity 28.9%; Pred. No. 8,7e-31;  
 Matches 165; Conservative 117; Mismatches 238; Indels 51; Gaps 14;

QY 10 TIFQAVLAVPAVITAMVSEFINILALVSPYLMQVYDRVLTSRNVSTLIVTVICVF 69  
 DB 145 TWFPATIKYRRIFETLVSVFLQFLALITPLEFQVYMDKVLVHRGFSTLVNTVIALSV 204  
 QY 70 LFLVGLLEALRTQVLVVRGKLFQVARDPIFKSVLDSTLS-----RKIGGQAFRDMQ 124  
 DB 205 VVVEIILSGLRITVPSHSTRIDVELGAKLFRHLALPIFENRNVGDVAVARELDQ 264

QY	125	VREFTGZ-GLIAFCDAWTPPEVSVNMLHPFECIALIIA--CIIFGLVAMNDNATKNP	181
Db	265	IRNFLTGCALTSVDLDLFEFTFEFVMMYSPKLIIVLLISLPCIAMSTI-----SP	317
QY	182	IQMTWASIAAQN-DAGSTLRNA---EYKAMGMGGLQARMRARDEQVANOQAASDA	236
Db	318	ILRRRLDEKFRANDNDOSFVIESASADITKLALVATPQMNINW---DKOLASVSA-DP	372
QY	237	GGAAVMS-----GIVFENIYOTLLGGGAVIATIDGKSAGAMTAGSTLVSRALAPLEGAV	291
Db	373	RYVYVLAITIGQGGVLDLQRTWMTIWMLGAIHVLISGDLSIQOLLTFNMLSGOVIAPVYRLA	432
QY	292	GOMKNYIGARGANDRLQTMJLRE--EKSDADHMPLEPPRGVLS-----AEAASILPQ	341
Db	433	QLMDQFOQYVIGISTRRLDVLNSPTENYQKLSLPEIGDLAFKHIRRYKRPDAPILLD--	490
QY	342	AQQTMMQASRIDAGAVALVYPSAAGKSSSLRGIVYWPVCAAGVIRLDGTYDKOMDPE	401
Db	491	-----DVMLSTVYQGEVYIGVYSGSGSKSTYLRKLQRFYIPENGQVLDIGHDLADPN	543
QY	402	KLGRHNVGLPDIDELFSGTVAONFARTEESOE--VTEATYTLAGVHMIOSLPAGYDTAI	460
Db	544	WLRQIGVYLDQDNLNKRSTRIDNIALTLDPKSMSPRYIYAKKLAGAMDPISELDEGTYTV	603
QY	461	GEGGASLISGGORLALARAFAFMFALLVLDEPNASLDQYGEVALMEAMRLKAAKRTYV	520
Db	604	GELGATSGGGORRIATARALVNNPRLLIFDEPATSALDYESEHIMQNKI--CHGRTVI	662
QY	521	FATHKVNLAQADYIMVINGVTSIDPEORR	551
Db	663	IIARLSTVYKADRIIVYMGKHLYEQQKHQ	693

ID	CYAB_BORPE	STANDARD	PRT	712 AA.
AC	P18770;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Cyclolysin secretion ATP-binding protein cyab.			
GN	CYAB.			
OS	Bordetella pertussis.			
OC	Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;			
CC	Bordetella.			
OX	NCBI_TaxID=520;			
RP	[1]			
RA	SEQUENCE FROM N.A.			
RX	STRAIN=18323;			
RT	MEDLINE=69091151; PubMed=2905265;			
RL	"Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;			
CC	"Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-			
CC	haemolysin bifunctional protein of Bordetella pertussis.";			
CC	EMBO J. 7:3997-4004(1988).			
CC	-1- FUNCTION: INVOLVED IN THE EXPORT OF CALMODULIN-SENSITIVE			
CC	ADENYLATE CYCLASE-HEMOLYSIN (CYCLOLYSIN).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. HLYB SUBFAMILY.			
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
DR	EMBL; X14199; CAA32412.1; -			
DR	PIR; S02386; BVBRCB.			
DR	InterPro; IPR003593; AAA.			
DR	InterPro; IPR001140; ABC_transporter_tmem.			
DR	InterPro; IPR003439; ABC_transportr.			
DR	InterPro; IPR001687; ATP_GTP_A.			

DR	Pfam: PF000664; ABC_membrane; 1.
DR	Pfam: PF000005; ABC_tran; 1.
DR	SMART; SM00382; AAA; 1.
DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW	Hemolysis; Transport; ATP-binding; Transmembrane.
FT	TRANSMEM 105 122 POTENTIAL.
FT	TRANSMEM 226 242 POTENTIAL.
FT	TRANSMEM 261 282 POTENTIAL.
FT	TRANSMEM 295 315 POTENTIAL.
FT	TRANSMEM 368 392 POTENTIAL.
FT	TRANSMEM 399 419 POTENTIAL.
FT	TRANSMEM 464 488 POTENTIAL.
FT	NP_BIND 505 512 ATP (BY SIMILARITY).
SQ	SEQUENCE 712 AA; 77969 MW; BA8DDC1450960675 CRC64;

Query Match	19.0%;	Score 533;	DB 1;	Length 712;
Best Local Similarity	27.0%;	Pred. No. 1.9e-29;		
Matches 166;	Conservative 105;	Mismatches 247;	Indels 96;	Gaps 15;

[illegible]

RESULT	12
ABCB6_HUMAN	
ABCB6_HUMAN	STANDARD;
Q9NP58; O75542;	PRT; 842 AA.
DT 16-OCT-2001 (Rel. 40, Created)	
DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	
ATP-binding cassette, sub-family B, member 6,	
	mitochondrial precursor



[illegible]



```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable ABC transporter ATP-binding protein Y4GM.
GN Y4GM.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Fiebelberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RT Perret X.;
RL "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000075; AAB91687.1; -.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 1.
DR SMART; SM00382; ABC_tran; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR Hypothetical protein; ATP-binding; Transport; Transmembrane; Plasmid.
KM TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 66 86 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT NP_BIND 379 386 ATP (POTENTIAL).
SQ SEQUENCE 586 AA; 64262 MW; 1095DFEB82620637 CRC64;

Query Match 15.9%; Score 445.5; DB 1; Length 586;
Best Local Similarity 25.1%; Pred. No. 1,7e-23;
Matches 143; Conservative 109; Mismatches 226; Indels 91; Gaps 15;

OY 33 INILAVSPYLMQVYDRVLTNRVSTLVTVICVFLVYGLLEALRQVLVVGGLKF 92
DB 39 VAVMTASAMIRDVNVSTVSKDIEKVFQVAVTVAIFAKGLATVYQSTFLSKAGNNI 98
OY 93 DGVARDFEKSVDLSTSRKIGIGQAFRMDQVFEPTGGLAFCDAPWTFVYIVSM 152
DB 99 IAHTRRLFEHVL-----RGLSFYSIYPSEELVRLTN-----AQAVRSVIELVVSFI 149
OY 153 HPFGILAIACIT-----IEGLAVMNDATK-NPIQVMTASIAQN 194
DB 150 RDLFSLGLAVVYIQPLISVSAVPGAILVTRKVRKIMELEIASIGQITQSV 209
OY 195 DAGSTLRNAYKAMGMGLOARWR-----ARRDEVAMQAASD-----AG 237
DB 210 QETST--GIRIVKAFLEDFRRRMDKYIGDVERANSIALLEAASSPIMETLSGPAIAG 267
OY 238 GAVWSGTRKVR-----NIVOTLLGGAY-----LAIDGKISAGAMAGSILYGRA 283
DB 268 VIALISGLVLDQNGTPELMSFTALL--AYEPARKLAMRISLSALGVMYQOLA 324
OY 284 LAPTEGAVGQKNYIGARGAMDRLQTMIREKKSADDDHPPCEPHGVSAEASITLPGAO 343
DB 325 DHRLE-----LTKNSA-----IPLDEGGEIRFKVNFSTYKNGE 359
OY 344 QPTMQASFRIDAGAAVALVPSAAGKSSLLRGIVGWPCAAVIRLDGYDKMDPEKL 403

```

```

DB 360 R-LFQNLVTFPPACKTTALVSPGAGKSSIIINILIRLTPDPVGSVTVDGHLKDPFRSL 418
OY 404 GRHGYLPQDLEFGSTVAONIAFTFEES-QEYIEATTLGVMEMQSLPMGVDFAIGE 462
DB 419 RDRIGFVGQDFLFSGTITKYNISLREGASDEEITIAKTNANDEFIMKPHGYDTEYGE 478
OY 463 GGASLSGGQORLALRAVFMPPALVDEPNASIDQGEVALMEAMRLKAARTVIFA 522
DB 479 NGIKLSGQKORITIRARMLRNALLFDEATSLDSESEIOIQALARL-TRKRTTITMI 537
OY 523 THKVNLAQADYIWINQVISPGEKDR 551
DB 538 AHRSTVTAADNIYMEGQVAEQPGGR 566

RESULT 14
YFIC_BACSU STANDARD; PRT; 604 AA.
AC P54719;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical ABC transporter ATP-binding protein yf1c.
GN yf1c.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96262713; PubMed=8704981;
RA Yamamoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;
RT "Determination of a 12 kb nucleotide sequence around the 76 degrees
RT region of the Bacillus subtilis chromosome.";
RL Microbiology 142:1417-1421(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50543; BAA09107.1; -.
DR EMBL; Z99108; CAB12651.1; -.
DR HSPSP; P13569; INBD.
DR SubList; BGI1850; yf1c.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR Hypothetical protein; ATP-binding; Transport; Transmembrane;
KM Complete proteome.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 172 192 POTENTIAL.
FT TRANSMEM 288 308 POTENTIAL.
FT TRANSMEM 510 530 POTENTIAL.
FT NP_BIND 399 406 ATP (POTENTIAL).
SQ SEQUENCE 604 AA; 67396 MW; 1998B4669A2E6133 CRC64;

Query Match 15.7%; Score 440.5; DB 1; Length 604;
Best Local Similarity 25.4%; Pred. No. 3.8e-23;
Matches 147; Conservative 113; Mismatches 232; Indels 87; Gaps 19;

```

Oy	16	VLVAPAVITAMWFSFIIINILATVSLPLMYOVDRVLTSSNGLVYIVYCEFLVYG	75
Db	50	ILVLMVAVISA-----IFGLGPYVIGKADHFIYGVKTVSGLLPVLILAIYIOS	101
Oy	76	LLEALRT--QVLYVGGLKFDVDARDPIFKFSYLDSTL-----SRKG-IGGAFRDMDOVRE	127
Db	102	LSLWFQNWMMITISQGVYER--MSESELFTHLHELPIPFDFKORHGLMSVTDIENVSS	159
Oy	128	FMVGGGLAFCDAPMTPEPVFL-VSMMLHFPFGILAIITICIIIFGLAVNNDNATKNPLDMAT	186
Db	160	TLNTSVIOLISLSSVTFEGVIAVMLYMSPP--LLTLILITLI-----PVMAAS	203
Oy	187	MASTA-----AONDG-----SLTRNAEYKRMGMGGLQARMPARDEQYAMQ	230
Db	204	LKWITINRFGKLFKEQOKKMLGDLNGLTIESVSGAKVIATYS-----REKQITAE	251
Oy	231	-----AASDAG-----GAVMSGIKYFRNIYOTLILGGGAYLAIDKISAGAMTAGSI	278
Db	252	FLEKNAALTKTSGFMNQITSGFIPRYMNSLNLSPETMAIAGLFPALKGWISIGSIYFAE	311
Oy	279	LVGRALAIIEGAVGMKVIYARGCAMDRLOMLREEKSADD-----HMPLEPPGVLSAE	333
Db	312	YSROFTRELNDLANQFNTMLSAIAGAEVFEVDLEKEDEKNAVHQPI--QTSIEFR	365
Oy	334	AASTLPPEAQPTMROASFRIDIAAVALVPSAAGKSSSLDRGIVGWPCAAGYIRLDGY	393
Db	370	DVSEFYDMGQO--TLKHLOFTYPAQOSIAFVPTGAKGTYYINLARRYEPKDGKILLDGT	428
Oy	394	DIKQMDPEKLRHVGYLEPQDILEFSGTVAAQNIARFTFEES--QEVIEATPLAGYHEMIQS	451
Db	429	DIKTTLFRASLTKRNMGFVLODSFLTFQGTIRENI--RYGRILASDOEVEAAATANAHSFIER	487
Oy	452	LPMGVDTAIGSGGASLSGGOSORLALAAVYRMALVILDEPNASLDPOYGVALMEAMKR	511
Db	488	LPKGDIVLTQSGSTISGGQKQLISIAAVLADPYLLIDTANTSIDPITVEVNIQELAR	547
Oy	512	LKAARKTYIFATHKVNLQAQDYIMVINOQYISDFGERD	550
Db	548	LMEGRTSVIIA-HRLNTIQRADQIYVLKNGMIEKGSMD	585
RESULT 15			
MSBA_ECOLI STANDARD; PRT: 582 AA.			
ID	MSBA_ECOLI	STANDARD;	PRT: 582 AA.
AC	P27299:		
DT	01-AUG-1992 (Rel. 23, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DT	01-MAR-2002 (Rel. 41, Last annotation update)		
DE	Probable transport ATP-binding protein msba.		
GN	MSBA OR B0914 OR Z1260 OR ECS0997.		
OS	Escherichia coli, and		
OS	Escherichia coli O157:H7.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Escherichia.		
OX	NCBI_TaxID=562, 83334;		
RA	NCBI_TaxID=562, 83334;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-K12:		
RC	MEDLINE-93172962; PubMed-8094880;		
RA	Katow M.L., Georgopoulos C.P.:		
RT	"The essential Escherichia coli msbA gene, a multicopy suppressor of		
RT	null mutations in the htrb gene, is related to the universal		
RT	conserved family of ATP-dependent translocators.";		
RL	Mol. Microbiol. 7:69-79(1993).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-K12 / MG1655;		
RA	MEDLINE-97426617; PubMed-9278503;		
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,		
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,		
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,		
RA	Mau B., Shao Y.;		

RT "The complete genome sequence of *Escherichia coli* K-12.";  
RL Science 277:1453-1474(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN-K12;  
RC MEDLINE=97061202; PubMed=8905232;  
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizubuchi K.,  
RA Mori H., Motomura K., Nakamura Y., Nishio H., Nishio Y., Saito N.,  
RA Samedji G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
RA Yano M., Horinouchi T.;  
RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome  
RT corresponding to the 12.7-28.0 min region on the linkage map.";  
RL DNA Res. 3:157-155(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX STRAIN-O157:H7 / ATCC 700927;  
RC MEDLINE=21074935; PubMed=11206551;  
RA Perina N.T., Plunkett G., Ili, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobbeck E.J., Davis N.W., Lim A., Dimlantia E.T., Potamouls K.,  
RA Apodaca J., Anantharaman T.S., Iain J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
RL Nature 409:529-533(2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX STRAIN-O157:H7 / RIMD 0509952;  
RC MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Chnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsuo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
RA Kufner S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RL DNA Res. 8:11-22(2001).  
RN [6]  
RP CHARACTERIZATION.  
RX MEDLINE=98241619; PubMed=9575204;  
RA Zhou Z., White K.A., Polissi A., Georgopoulos C., Raetz C.R.;  
RT "Function of *Escherichia coli* MSBA, an essential ABC family  
RT transporter, in lipid A and phospholipid biosynthesis.";  
RL J. Biol. Chem. 273:12466-12475(1998).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (4.5 ANGSTROMS).  
RX MEDLINE=21432170; PubMed=11546864;  
RA Chang G., Roth C.B.;  
RT "Structure of MSBA from *E. coli*: a homolog of the multidrug resistance  
RT ATP binding cassette (ABC) transporters.";  
RL Science 293:1793-1800(2001).  
RN [8]  
RP FUNCTION: INVOLVED IN THE SECRETION OF LIPID A AND POSSIBLY ALSO  
RC GLYCEROPHOSPHOLIPID TRANSPORT.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z11796; CAA77839.1; -;  
DR EMBL: AE000193; AAC74000.1; -;  
DR EMBL: D90729; BA35658.1; -;  
DR EMBL: D90730; BA35660.1; -;  
DR EMBL: AE005281; AAG55399.1; -;  
DR EMBL: AP002553; BAB34420.1; -;  
DR EMBL: S21588; S21588;  
DR PIR: S27996; S27998;  
DR PIR: S27996; S27998.

Search completed: July 19, 2002, 10:49:04  
Job time: 404 sec

DR	PDB: 1JSO: 12-SEP-01.
DR	EcoGene; BG10613: msbA.
DR	InterPro; IPR003593: AAA.
DR	InterPro; IPR001140: ABC_transporter_tmem.
DR	InterPro; IPR003439: ABC_transportr.
DR	InterPro; IPR001687: ATP_GTP_A.
DR	Pfam; PF00664: ABC_membrane1.
DR	Pfam; PF00005: ABC_tran; 1.
DR	SMART; SM00382: AAA; 1.
DR	PROSITE; PS00211: ABC_TRANSPORTER; 1.
KW	ATP-binding; Transport; Inner membrane; Transmembrane; 3D-structure;
KW	Complete proteome.
FT	DOMAIN 1 21 CYTOPLASMIC.
FT	TRANSMEM 22 52
FT	DOMAIN 53 64 EXTRACELLULAR.
FT	TRANSMEM 65 96
FT	DOMAIN 97 139 CYTOPLASMIC, INTRACELLULAR DOMAIN 1.
FT	TRANSMEM 140 164
FT	DOMAIN 165 167 EXTRACELLULAR.
FT	TRANSMEM 168 192
FT	DOMAIN 193 252 CYTOPLASMIC, INTRACELLULAR DOMAIN 2.
FT	TRANSMEM 253 272
FT	DOMAIN 273 280 EXTRACELLULAR.
FT	TRANSMEM 281 301
FT	DOMAIN 302 582
FT	NP_BIND 376 383
FT	CYTOPLASMIC, INTRACELLULAR DOMAIN 3.
FT	ATP.
QO	SEQUENCE 582 AA; 64460 MW; 3D52A348217A47EE CRC64;

Query Match	15.7%	Score 440	DB 1	length 582
Best Local Similarity	26.7%	Pred. No. 4e-23		
Matches 152	Conservative 112	Mismatches 248	Indels 58	Gaps 17

[illegible]

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:41:45 ; Search time 53.06 Seconds  
(without alignments)  
1809.502 Million cell updates/sec

Title: US-09-913-414-4

Perfect score: 2807

Sequence: 1 MFKRSQAKPTFDQAVLVAR.....MVINQGVISDFGERDRCWPS 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rviro:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2765	98.5	578	16	085350
2	2759	98.3	578	2	09RMNO
3	1206.5	43.0	583	2	09Z3C5
4	1195	42.6	580	2	09RHR2
5	1193.5	42.5	583	2	09ZNI9
6	1184	42.2	578	2	09ZG94
7	1182.5	42.1	578	2	054456
8	1177.5	41.9	579	2	067993
9	1176.5	41.9	588	2	09XB64
10	1172	41.8	583	2	087808
11	1164.5	41.5	599	2	09KGS6
12	1156	41.2	600	16	09HYJ8
13	1138	40.5	617	16	054416
14	1088.5	38.8	617	16	092MX3
15	1077	38.4	582	16	033678
16	1068	38.0	581	16	098L68

17	1060.5	37.8	570	16	0988E6	0988E6 rhizobium l
18	1052.5	37.5	571	2	09X6N6	09X6N6 rhizobium l
19	1044.5	37.2	564	16	0677184	0677184 aquilex aeo
20	1039	37.0	576	2	053368	053368 serratia ma
21	1038	37.0	586	2	054451	054451 serratia ma
22	1038	37.0	570	2	057040	057040 rhizobium l
23	1036.5	36.9	579	2	085375	085375 proteus mlr
24	1033	36.8	570	2	005693	005693 rhizobium l
25	1012	36.1	583	16	09ZDL3	09ZDL3 rickettsia
26	1011	36.0	584	16	09Z1J1	09Z1J1 rickettsia
27	958	34.1	589	2	P66439	P66439 rhizobium m
28	958	34.1	589	16	092V33	092V33 rhizobium m
29	957	34.1	589	2	09X7L7	09X7L7 rhizobium m
30	913	32.5	712	16	092LS9	092LS9 rhizobium m
31	892	31.8	581	16	098159	098159 rhizobium m
32	596.5	21.3	706	2	047462	047462 escherichia
33	592.5	21.1	706	2	046717	046717 escherichia
34	590.5	21.0	705	2	09LC57	09LC57 escherichia
35	582.5	20.8	708	2	093FG6	093FG6 pasteurella
36	581.5	20.7	708	2	093FH0	093FH0 pasteurella
37	580.5	20.7	708	2	093FH3	093FH3 pasteurella
38	580.5	20.7	708	2	0934E0	0934E0 pasteurella
39	580.5	20.7	708	2	0934A3	0934A3 pasteurella
40	577.5	20.6	708	2	093FG8	093FG8 pasteurella
41	577.5	20.6	708	2	093FG4	093FG4 mannheimia
42	577.5	20.6	708	2	093FG2	093FG2 mannheimia
43	577.5	20.6	708	2	0933I3	0933I3 mannheimia
44	577.5	20.6	1011	16	P74176	P74176 synechocyst
45	576.5	20.5	708	2	093FH6	093FH6 pasteurella

## ALIGNMENTS

RESULT 1  
ID 085350 PRELIMINARY: PRT: 578 AA.  
AC 085350:  
DT 01-NOV-1998 (TREMBLrel. 08. Created)  
DT 01-MAY-2000 (TREMBLrel. 13. Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19. Last annotation update)  
DE ABC TRANSPORTER (RSAA SECRETION SYSTEM, ATP-BINDING PROTEIN RSAD).  
GN RSAD OR CC1008.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_TaxID=69394;  
RN [1]  
RP SEQUENCE OF 1-98 FROM N.A.  
RC STRAIN-ATCC 19089 / CB15, AND NA1000;  
RX MEDLINE=85054630; PubMed=6209263;  
RA Smit J.K., Agabian N.;  
RT "Cloning of the major protein of the Caulobacter crescentus periodic surface layer: detection and characterization of the cloned peptide by protein expression assays."  
RT J. Bacteriol. 180:1137-1145(1984).  
RL [2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19089 / CB15, AND NA1000;  
RX MEDLINE=98292737; PubMed=9620954;  
RA Avram P., Smit J.;  
RT "The Caulobacter crescentus paracrystalline S-layer protein is secreted by an ABC transporter (Type I) secretion apparatus."  
RT J. Bacteriol. 180:3062-3069(1998).  
RL [3]  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19089 / CB15, AND NA1000;  
RA Avram P., Smit J.K.;  
RT submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
RL [4]  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19089 / CB15, AND NA1000;

RA Avram P.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC SRRAIN-ATCC 19089 / CB15;  
RX MEDLINE-21173698; PubMed-11259647;  
RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,  
Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
Uterback T., Tran K., Wolf A., Yamathayan J., Ermolaeva M., White O.,  
Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,  
RT "Complete genome sequence of *Caulobacter crescentus*.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
(ABC TRANSPORTERS).  
CC EMBL: AF062345; AAC38666.2; -;  
DR EMBL: AE005779; AAC22992.1; -;  
DR HSSP: P13569; INBD.  
DR TIGR: CC1008; -;  
DR InterPro: IPR001140; ABC\_transporter\_tmam.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 1.  
DR Pfam: PF00005; ABC\_tran; 1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
KW ATP-binding; Complete proteome; Transport.  
SQ SEQUENCE 578 AA; 62019 MW; 1E53C2DA497675C5 CRC64;

Query Match 98.5%; Score 2765; DB 16; Length 578;  
Best Local Similarity 99.8%; Pred. No. 4.9e-160;  
Matches 549; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MFKRSAGPTTFDQAVLVARPAVITAMVFSFPIINITALVSPLYMLOYDRLVTSRNVSTL 60  
DB 1 MFKRSAGPTTFDQAVLVARPAVITAMVFSFPIINITALVSPLYMLOYDRLVTSRNVSTL 60  
OY 61 IYLTVICVFLVYGLLELRTQVLVGRGLKPDGVARDPFEKSVSDSTLSRKIGIGQAFR 120  
DB 61 IYLTVICVFLVYGLLELRTQVLVGRGLKPDGVARDPFEKSVSDSTLSRKIGIGQAFR 120  
OY 121 DMDQVREFMTGGLIAFCDAFWTPVFEIVYSWMLHPFFGILAIITACIIIFGLAVMNDNATKN 180  
DB 121 DMDQVREFMTGGLIAFCDAFWTPVFEIVYSWMLHPFFGILAIITACIIIFGLAVMNDNATKN 180  
OY 121 DMDQVREFMTGGLIAFCDAFWTPVFEIVYSWMLHPFFGILAIITACIIIFGLAVMNDNATKN 180  
DB 121 DMDQVREFMTGGLIAFCDAFWTPVFEIVYSWMLHPFFGILAIITACIIIFGLAVMNDNATKN 180  
OY 181 PIQMATMASIAQNDAGSTLRNAEYKAMGMGGLQARWRARDEQVAAQAAASDAGAV 240  
DB 181 PIQMATMASIAQNDAGSTLRNAEYKAMGMGGLQARWRARDEQVAAQAAASDAGAV 240  
OY 241 MSGIKVFRNIYVOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300  
DB 241 MSGIKVFRNIYVOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300  
OY 301 RCAMDRLQTMLEKESADDDHMLPEPRGVLSAEASILPPGAQOQPTMRQASFRIDAGAAV 360  
DB 301 RCAMDRLQTMLEKESADDDHMLPEPRGVLSAEASILPPGAQOQPTMRQASFRIDAGAAV 360  
OY 361 ALVGPAAKSSSLIRGIVGWPCAGVIRLDGYDIKQMPDEKIGRHVGLPDIIEFSGT 420  
DB 361 ALVGPAAKSSSLIRGIVGWPCAGVIRLDGYDIKQMPDEKIGRHVGLPDIIEFSGT 420  
OY 421 VAONIRFTEFESQEVIEATLAGVHEMIQSLPMGYDTAIGEGASLSGGORRLALARA 480  
DB 421 VAONIRFTEFESQEVIEATLAGVHEMIQSLPMGYDTAIGEGASLSGGORRLALARA 480  
OY 481 VFRMPALVLVDEPNASLDQVGEVALMEAMKRLKAARKTVIFATHKYNLLAQADYIMVINO 540  
DB 481 VFRMPALVLVDEPNASLDQVGEVALMEAMKRLKAARKTVIFATHKYNLLAQADYIMVINO 540  
OY 541 GVISDFGERD 550  
DB 541 GVISDFGERD 550

DB 541 GVISDFGERD 550

RESULT 2  
O99RMO PRELIMINARY; PRT; 578 AA.  
AC O99RMO;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE ABC\_TRANSPORTER.  
GN RSAD.  
OS *Caulobacter crescentus*.  
OC Bacteria; Proteobacteria; alpha subdivision; *Caulobacter* group;  
OC *Caulobacter*.  
OX NCBI\_TaxID=69394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SRRAIN-J54000;  
RA Bingle W.H., Avram P.A., Nornellini J.F., Smit J.R.;  
RT "The secretion signal of the *Caulobacter crescentus* S-layer protein is  
located within the C-terminal 82 amino acids of the molecule.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE ABC\_TRANSPORTER FAMILY.  
DR EMBL: AF193064; AAF07962.1; -;  
DR HSSP: P13569; INBD.  
DR InterPro: IPR001140; ABC\_transporter\_tmam.  
DR InterPro: IPR003439; ABC\_transporter.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane; 1.  
DR Pfam: PF00005; ABC\_tran; 1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
KW ATP-binding; Transport.  
SQ SEQUENCE 578 AA; 62006 MW; 8661C449E0A47ADB CRC64;

Query Match 98.3%; Score 2759; DB 2; Length 578;  
Best Local Similarity 99.6%; Pred. No. 1.1e-159;  
Matches 548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MFKRSAGPTTFDQAVLVARPAVITAMVFSFPIINITALVSPLYMLOYDRLVTSRNVSTL 60  
DB 1 MFKRSAGPTTFDQAVLVARPAVITAMVFSFPIINITALVSPLYMLOYDRLVTSRNVSTL 60  
OY 61 IYLTVICVFLVYGLLELRTQVLVGRGLKPDGVARDPFEKSVSDSTLSRKIGIGQAFR 120  
DB 61 IYLTVICVFLVYGLLELRTQVLVGRGLKPDGVARDPFEKSVSDSTLSRKIGIGQAFR 120  
OY 121 DMDQVREFMTGGLIAFCDAFWTPVFEIVYSWMLHPFFGILAIITACIIIFGLAVMNDNATKN 180  
DB 121 DMDQVREFMTGGLIAFCDAFWTPVFEIVYSWMLHPFFGILAIITACIIIFGLAVMNDNATKN 180  
OY 121 DMDQVREFMTGGLIAFCDAFWTPVFEIVYSWMLHPFFGILAIITACIIIFGLAVMNDNATKN 180  
DB 121 DMDQVREFMTGGLIAFCDAFWTPVFEIVYSWMLHPFFGILAIITACIIIFGLAVMNDNATKN 180  
OY 181 PIQMATMASIAQNDAGSTLRNAEYKAMGMGGLQARWRARDEQVAAQAAASDAGAV 240  
DB 181 PIQMATMASIAQNDAGSTLRNAEYKAMGMGGLQARWRARDEQVAAQAAASDAGAV 240  
OY 241 MSGIKVFRNIYVOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300  
DB 241 MSGIKVFRNIYVOTLLIGGAYLAIDCKISAGAMINGSLVGRALPIEGAVGOMKNYIGA 300  
OY 301 RCAMDRLQTMLEKESADDDHMLPEPRGVLSAEASILPPGAQOQPTMRQASFRIDAGAAV 360  
DB 301 RCAMDRLQTMLEKESADDDHMLPEPRGVLSAEASILPPGAQOQPTMRQASFRIDAGAAV 360  
OY 361 ALVGPAAKSSSLIRGIVGWPCAGVIRLDGYDIKQMPDEKIGRHVGLPDIIEFSGT 420  
DB 361 ALVGPAAKSSSLIRGIVGWPCAGVIRLDGYDIKQMPDEKIGRHVGLPDIIEFSGT 420  
OY 421 VAONIRFTEFESQEVIEATLAGVHEMIQSLPMGYDTAIGEGASLSGGORRLALARA 480  
DB 421 VAONIRFTEFESQEVIEATLAGVHEMIQSLPMGYDTAIGEGASLSGGORRLALARA 480  
OY 481 VFRMPALVLVDEPNASLDQVGEVALMEAMKRLKAARKTVIFATHKYNLLAQADYIMVINO 540  
DB 481 VFRMPALVLVDEPNASLDQVGEVALMEAMKRLKAARKTVIFATHKYNLLAQADYIMVINO 540

```
|||||
Db 481 VFRPALVDEPNASLDQGEVALMEAMKRLKAARKTVFATHKVNLLAQADYIMVING 540
OY 541 GVISDFGERD 550
Db 541 GVISDFGERD 550

RESULT 3
O923C5 PRELIMINARY: PRT; 583 AA.
AC O923C5;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE SAPP.
GN SAPP.
OS Campylobacter fetus.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=196;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=84-107, AND 23D;
RX MEDLINE=99069317; PubMed=9851986;
RA Thompson S.A., Shedd O.L., Ray K.C., Belns M.H., Jorgensen J.P.,
  Blaser M.J.;
RT "Campylobacter fetus surface layer proteins are transported by a type
  I secretion system.";
RT J. Bacteriol. 180:6450-6458(1998).
CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AF071883; AAC97197.1; -
DR EMBL: AF027405; AAC97164.1; -
DR HSPF: P13569; INBD.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF00664; ABC_membrane.1.
DR Pfam: PF00005; ABC_tran.1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Transport.
SQ SEQUENCE 583 AA; 63958 MW; 891C8046136B7FAB CRC64;

Query Match 43.0%; Score 1206.5; DB 2; Length 583;
Best Local Similarity 45.5%; Pred. No. 1.9e-65;
Matches 245; Conservative 110; Mismatches 181; Indels 3; Gaps 3;

OY 14 QAVLVARPAVITAMVRSFINITLALVSPLYMLQVDRVLTSRVSTLIVTCVFLFLV 73
Db 13 ETIKSKNCIVIAIRFSFVNLMILPPLMLQVDRVVTSRSESTLFLTLVLFLEFT 72
OY 74 YGLLEALRTQVLRGGLKFDGVARDPIFKSVLDSTLSRKG-IGQAFRMDQVREFM-TG 131
Db 73 MAIFELRSRLILVEFNOQDIDNTDRITYDAIFKLARSYGRVTSQMGDNLAKQVNSTN 132
OY 132 GLIACDPAFWTPVEVIVSMNLHPFEGILAIACIIIFGLAVNMNDATKNPDIOMATNASTA 191
Db 133 GJFAFADAWMLPYIYAILTFVHFWMGFIAAAIILFVALLNEKATKDKLKSNDPTYKN 192
OY 192 AQNDASTSTRNNEVMKAMMGGLQARRRAREQVAMQAAASDAGAVMSGIKVFRNIV 251
Db 193 EMRLIDMNLNRSNVEVINAMGNNNLTKIKNKHHTFNLNSHSDASIKAGVYTNIKATRVTS 252
OY 252 QTLILGGAGVLAIDKRTISAGAMIAGSILVGRALAPIEGAVGQKNYIGARGMDRLQTM 311
Db 253 QSMGLGLAGVIVLAKMEVSGMMIAGSILGRALAPIDILASKSVKNTKRESEERDKFL 312
OY 312 REKSKADHMLPLPERPGVISAESAISLPEGAOQPTMKQASFRIDAGAAVALVPSAAGS 371
Db 313 HDPPEVKDKLSLDPDPRGDIACEAISLIPPSAKQPSLIGVSFALDADMCALIGPSAAGS 372
```

```
OY 372 SLIRGIVWPCAGVIRLDGYDIKQWDPKLGHRHGYLPQDIELEFSGVIAQNIARFTEE 431
Db 373 SLARALIGIMPAHGVAVRVDCADINQYSDALGFVGYLQDVELFPGTIAENIARGEL 432
OY 432 ESEQIEATLAVGHENIQLPMGYDTAIGEGASLSGGQRQRLATARAVERMPALIVD 491
Db 433 DSAVAVEAAKSAVNHMILRLPDGYDTKIGLGMSTLSGGQRQRLALARKKPKIIVLD 492
OY 492 EPNASLDQGEVALMEAMKRLKAARKTVFATHKVNLLAQADYIMVINGVIDDFGERD 550
Db 493 EPNASLDEEGERLVAALLAMK-GKATITLITRKNLVLAQVADKIAVLNAGOLVYFGERD 550

RESULT 4
O9RHT2 PRELIMINARY: PRT; 580 AA.
AC O9RHT2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ABC PROTEIN.
GN HASD.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NO.33;
RX MEDLINE=20069636; PubMed=10601212;
RA Idel A., Kawai E., Aketsuka H., Omori K.;
RT "Cloning and characterization of the Pseudomonas fluorescens ATP-
  binding cassette exporter, hasDEF, for the heme acquisition protein
  hasA.";
RT J. Bacteriol. 181:7545-7551(1999).
CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AB023289; BAA8492.1; -
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC_transporter_tmam.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00664; ABC_membrane.1.
DR Pfam: PF00005; ABC_tran.1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Transport.
SQ SEQUENCE 580 AA; 61803 MW; 3A6B3696E3798B3C CRC64;

Query Match 42.6%; Score 1195; DB 2; Length 580;
Best Local Similarity 45.5%; Pred. No. 9.3e-65;
Matches 240; Conservative 106; Mismatches 176; Indels 6; Gaps 3;

OY 28 VFEFFINITLALVSPLYMLQVDRVLTSRVSTLIVTCVFLFLVYGLLEALRTQVLR 87
Db 24 LFSAVINLMLALVALVLMQVYDRVLASGNMTLLMTLMLGLFGLMGALEWRSQVYIR 83
OY 88 GLKFDGVARDPIFKSVLDSTLSRKG--IGQAFRMDQVREFMTG-GLIACDPAFWTP 143
Db 84 LGTQMDRLNQRVYDAFAEQL--KGTQAAQOALSDLTLLRQFATQOALFAFDDAWFP 141
OY 144 VEVIVSMNLHPFEGILAIACIIIFGLAVNMNDATKNPDIOMATNASTAONDAGSTIRNA 203
Db 142 VYLLVLFLEFHPWGLVALACALMLMALVNMHICQAPLQAOQLSISASQOATANLRNA 201
OY 204 EVKAKGMGMGLOARRRAREDEVAMQAAASDAGAVMSGIKVFRNIVOTLILGGAGVLA 263
Db 202 EALEAMGLATLARFAQOAFLAQNLASSEKTAIISMSKVRALQSLVIGLGLALLA 261
OY 264 IDKISAGAMIAGSILVGRALAPIEGAVGQKNYIGARGMDRLQTMREKSGADHMLPL 323
Db 262 VQCAITPGMIAGSILMGVLSPLDQIGVWKQWSSARLAYERLITMLANPARTERMSL 321
```

Oy	324	PERGVLISAAASLLEDPGAAQPTMRQASPRIDAGAVALVYGSAAKSKSLLRGVWPC	383
Db	322	PASGQLTVEQVSCACAPGRRPALANLWFSLEPAGDVLGYIGSGCGKSLARLVGAMTP	381
Oy	384	AAGVTRLDGYDIKQWDPBEKLGRRHVGYPDLIELFSGTVAQNIAITFEFSEOEVIATLA	443
Db	382	MAGVRLRDGADLAAQKQOLGPHIGTLPDQIOLFAGSIADNTAIFEDYPSDKVLAQAOMA	441
Oy	444	GVHMIQSLPMGVDTALGEGSASLSGGQQRQRLALRAVPRMALLVDEPNASIDQCEV	503
Db	442	GVHQLILQLPQGDYDTQIGEGSAGLSGGQQRQVLAALNATYIGLPALLVDEPNSLDEAGEQ	501
Oy	504	ALPMAMRLKAARKTVEFATFKNNLAAQADYIMVINOVIYSGEDR	551
Db	502	ALLAATLTLKAQRRTLVLTHTKPTVALTDQLLILRDQLQAFGPTAR	549
RESULT	5		
O9ZN19	O9ZN19	PRELIMINARY;	PRT: 583 AA.
AC	O9ZN19		
DT	01-MAY-1999 (TREMBLrel. 10, Created)		
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	ABC PROTEIN.		
GN	APRDPF33.		
OS	Pseudomonas fluorescens.		
OC	Bacteria: Proteobacteria: gamma subdivision; Pseudomonadaceae;		
OC	Pseudomonas		
OX	NCBI_TaxID=294;		
RN	11		
RN	SEQUENCE FROM N.A.		
RP	STRAIN-NO.33;		
RC	MEDLINE=99453738; Pubmed=10524213;		
RX	Kawai E., Idei A., Kimura H., Shinazaki K., Akatsuka H., Omori K.;		
RA	"The ABC-exporter genes involved in the lipase secretion are clustered		
RT	with the genes for lipase, alkaline protease, and serine protease		
RT	homologues in Pseudomonas fluorescens no. 33."		
RL	1-biochim. Biophys. Acta 1446:377-382(1999).		
CC	-1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.		
DR	EMBL: AB015053; BAA36463.1; -		
DR	MEROPS: C39.0PM; -		
DR	InterPro: IPR003593; AAA.		
DR	InterPro: IPR001140; ABC_transporter_tmam.		
DR	InterPro: IPR003439; ABC_transport.		
DR	InterPro: IPR001687; ATP_GTP_A.		
DR	Pfam: PF00664; ABC_membrane; 1.		
DR	Pfam: PF00005; ABC_tran; 1.		
DR	SMART: SM00382; AAA. 1.		
DR	SMART: PS00211; ABC_TRANSPORTER; 1.		
DR	ATP-binding; Transprot.		
SO	SEQUENCE 583 AA: 62919 MW: 153D27B628E19B8B CRC64:		

Query Match	42.5%	Score 1193.5	DB 2	Length 583
Best Local Similarity	44.4%	Pred. No. 1.2e-64		
Matches 244	Conservative 112	Mismatches 180	Indels 13	Gaps 3
QY	1	MFRSGAKPTIIDQAVLVARPAVITAMVFSFNNIALATSPYMLQYDVRVLTNRVSTL	60	
Db	10	LKRALGDYKSI-----LISGCFETALINVLMLPVSIMYLQYDVRVLSQNETLL	58	
QY	61	IYLVIVICVFLFYGLLELRLQVLRGSLKRTDGAARDPIFESVLDSITSLR-KGIGGQAF	119	
	::: :: ::	::: :: ::	::: :: ::	::: :: ::
	::: :: ::	::: :: ::	::: :: ::	::: :: ::
Db	59	VMLTLVWVFEEFAGIELEIVRSFIVIRIGISQLERRNRLNLYVAAAFERNLQREBGHGGSL	118	
QY	120	RMDDQREEMTG-GLYAFCDAPMTPEVIVSMNLHPHFGLIITCIIIFGLAVNMNDAT	178	
	::: :: ::	::: :: ::	::: :: ::	::: :: ::
	::: :: ::	::: :: ::	::: :: ::	::: :: ::
Db	119	GDLTHIRKQITTGALFAAFDAPREPIFLYVITFLFNWILGVLASAGAVLIGLACLNEYLT	178	
QY	179	KNPIDQATWASIAQNDAGSTLRNAEVMKAMGMGGGLQARWRARRDEQVWQAAAAADAG	238	
Db	179	KKPLGASGYSQOSTQATLSHLNNAITQALMGKGLRRKRWVHSQFGLGNKKASDTGS	238	

OY	239	A V M S G I K Y F R N I V T L L G G A Y I A I D C K I S A G A M I A G S I L V E R A L A P I E G A Q O G K N I	298
Db	239	I I S S I S K S L R I C I O S I V L G I C A L L Y I G D M T A G M I A G S I I M K R I S P I D O L L A V K W S	298
OY	299	G A R G A M P B L O T M L E E K S A D D H M P L P E P R G V L S A E A S I L P P C A O O P T P M R O A S F R I D A G A	358
Db	299	S A K I A Y O R L D N I M K E F F P P O G E O M A L P A P K G V S F E O V S A C P P G R R P Y T L I Q V S F N I A G E	358
OY	359	A V A L V E S P A S K S S I L K G I V S W P C A G A V I R L D G Y D I K O M D P E K I R N H V G Y L P O D I E L S	418
Db	359	V L G V I G A S G S K S T L A R V I A V W V P T L A G Y R L D S A D I H R M D R D D I G P H I G I L P Q D I E L S	418
OY	419	G T V A O N I A R F T E F E S Q E V I E A T L A G V H E M I O S L P M G Y D T A I G B G A S I S G G O R O L A T A	478
Db	419	G S I A D N I N R F E D A D P Q U I V A A O A G V H E L I R L P H G Y T P V L D E G S G I S G O K O R A I L A	478
OY	479	R A V F R M P A L V L D E P N A S L D O V E G A V L M E A M K R L K A K R V I F A T H K V N L L A O A D I N V I	538
Db	479	R A L V G G P R L I V L D E P N S M I D T V G E A A L A S A I M O K K A G S T V I V L T H R S S A L A O A D K L I V L	538
OY	539	N O G V I S D E G	547
Db	539	N E G R L Q A F G	547

```

RESULT      6
ID           092G94      PRELIMINARY;      PRT;      578 AA.
AC           092G94;
DT           01-MAY-1999 (TREMBLrel. 10, Created)
DT           01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT           01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE           ABC_TRANSPORTER TLID.
GN           TLID.
OS           Pseudomonas fluorescens.
OC           Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC           Pseudomonas.
OX           NCBI_TaxID=294;
RN           [1]
RP           SEQUENCE FROM N.A.
RC           STRAIN-SIK W1.
RX           MEDLINE=99173908; PubMed=10074078;
RA           Ahn J.H., Pan J.G., Rhee J.S.;
RT           "Identification of the tlidEF ABC transporter specific for lipase in
RT           Pseudomonas fluorescens SIK W1."
RL           J. Bacteriol. 181:1847-1852(1999).
CC           -1. SIMILARITY: BELONGS TO THE ABC_TRANSPORTER FAMILY.
DR           EMBL, AF083061, AAD09853.1; -.
DR           InterPro: IPR003593; AAA.
DR           InterPro: IPR001140; ABC_transporter_tmam.
DR           InterPro: IPR003439; ABC_transportr.
DR           InterPro: IPR01687; ATP_GMP_A.
DR           Pfam: PF00664; ABC_membrane.1.
DR           Pfam: PF00005; ABC_tran.1.
DR           SMART; SM00382; AAA.1.
DR           PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR           ATP-binding; Transport.
SQ           SEQUENCE 578 AA; 62561 MW; 9EF2DED29432D697 CRC64;

```

Query Match	42.28;	Score 1184;	DB 2;	Length 578;
Best Local Similarity	44.88;	Pred. No. 4,36-64;		
Matches 243;	Conservative 106;	Mismatches 192;	Indels 2;	Gaps 2
QY	12	PQDAVVARPAVITAMVESFNTIILALVSPMLQYDVRILSRNNSTLIILVILVICVLEF	71	
		: : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :		
Db	20	FLQGVGEYSILISVGCFTALINLMLVPSIYMLQYDVRILSRNNSTLIILVILVICVLEF	79	
		: : :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :		
QY	72	LVYGLLEALRTQVIVAGGLKFDGVADPLFEKSVLDSLSLR-KGIGGQARPDMDQVREPM	130	
		: : :   :   : : :   : : :   :   :   :   :   :   :   :   :   :   :		
Db	80	AFIGTLEIRSFVILVIGSOLERFNLRYVKAFFENLRGGCHAGQADLTLLNLFOT	139	



**QY** 131 G-GAFDAPPTPEVIVSMILHFEFILLAIACIIIFGLAVMMDNATKPIQMATAAS 189  
| | | | | :  
**Dd** 140 GALFAFDPAEPFLYLVLITLLENWLVGLATAGAVLLIGLACLNEIYLRKPLGEGAYS 199  
| | | | | :  
**QY** 190 IAAQNDA GSTLRNMEVMKMGWGGLQARMRARBDEOVAMQAASDAGAVMSGIKVERN 249  
| | | | | :  
**Dd** 200 QSSQLTASHLHNMTAQMGMLGARLRKRFVAHSOFLGLQNTASTDSGVTSLSKTIRL 259  
| | | | | :  
**QY** 250 IVOTLLLGGAAYLAIDGKISAGAMTAGSTLVGRALAPIEGAVGOKNITGANGANDRLQT 309  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**Dd** 260 CLQSIVLIGLALVLTGKGMTAGMIAGSILMKRVSLPSIDOLLTAWKWSAALAAQRDLDD 319  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**QY** 310 MLRREKSADHDHPLEDPREGVLSAEASAILPPCAOPTRMQASFRIIDACAVALVGPSAG 369  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**Dd** 320 LLREPPDSEPMKLPAPIRGVYSFEQVSAGPPRRRTPLHOVSFTLAGAVGLGASSGS 379  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**QY** 370 KSSLRLIGVGPWCAGAYIRLDGYDIKWMDPEKLRHVGYLPQDIELFSGTVAQNARTP 429  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**Dd** 380 KSTLARVLVGVWPTLGCTVRLDGADIHRWDREDLTPHGILYLPQDIIELESGSTADNIARR 439  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**QY** 430 EEESGEVEATNLTVGHHEMIOSLPNGCYDTATGEGASISGGOROLALARAERFMALLV 489  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**Dd** 440 QADPALVYAQAQAKQVHLIRLPLEGYDTLGDNSGGISGGKKORALARALTYGGRLIV 499  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**QY** 490 LDEPNASLDIOYEVALAMEAMRKLRKAARKRTYLFATRKNVLLAADYIMVINGVISDFGER 549  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**Dd** 500 LDEPNNDIPVGEALAGATVOMKQGSSVYLVTFRSSALAQADRLVLYNEGRLQLARA 559  
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :  
**QY** 550 DRC 552  
||  
**Dd** 560 RRC 562  
||  
**RESULT** 7  
**ID** 054456 PRELIMITARY; PRT: 588 AA.  
**AC** 054456:  
**DT** 01-NOV-1996 (TREMBLrel\_01, Created)  
**DT** 01-NOV-1996 (TREMBLrel\_01, Last sequence update)  
**DT** 01-DEC-2001 (TREMBLrel\_19, last annotation update)  
**DE** lipB.  
**GN** lipB.  
**OS** Serratia marcescens.  
**OC** Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
**CC** Serratia.  
**OX** NCBI\_TaxID=615;  
**RN** [1]  
**RP** SEQUENCE FROM N.A.  
**RC** STRAIN-SR41;  
**RX** MEDLINE-96062219; PubMed-7592412;  
**RA** Akutsuka H., Kawai E., Omori K., Shibatani T.;  
**RT** "The three genes lipB, lipC and lipD involved in the extracellular  
**RT** secretion of the serratia marcescens lipase which lacks an N-terminal  
**RT** signal peptide.";  
**RL** J. Bacteriol. 177:6381-6389(1995).  
**CC** -I- SIMILIARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
**DR** EMBL; D49826; BAA08631.1; -.  
**DR** HSSP; P13569; INBD.  
**DR** InterPro; IPR003593; AAA.  
**DR** InterPro; IPR001140; ABC\_transporter\_tmem.  
**DR** InterPro; IPR003439; ABC\_transportr.  
**DR** InterPro; IPR001687; ATP\_GTP\_A.  
**DR** Pfam; PF00664; ABC\_membrane; 1.  
**DR** Pfam; PF00005; ABC\_tran; 1.  
**DR** SMART; SM00382; AAA; 1.  
**DR** PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
**FW** ATP-binding; transport.  
**SEQUENCE** 588 AA; 64205 MW; 4FD37DAB30A0EC75 CRC64;

Query Match	42.1%;	Score 1182.5;	DB 2;	Length 588;
Best Local Similarity	44.9%;	Pred. No. 5.4e-64;		
Matches 236;	Conservative 110;	Mismatches 177;	Indels 3;	Gaps 3;

[illegible]

RESULT	8		
ID	067993	PRELIMINARY;	PRT; 579 AA.
AC	067993;		
DT	01-AUG-1998 (TREMBLrel. 07, Created)		
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	ZINC-PROTEASE TRANSPORTER.		
GN	APRD.		
OS	Pseudomonas fluorescens.		
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;		
CC	Pseudomonas		
OX	NCBI_TaxID=294;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CY091.		
RX	MEDLINE=98162052; PubMed=9501431;		
RA	Liao C.H., McCallus D.E.;		
RT	"Biochemical and genetic characterization of an extracellular protease		
RL	from Pseudomonas fluorescens CY091.";		
CC	Appl. Environ. Microbiol. 64:914-921(1998).		
CC	-I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.		
DR	HMBL: AF004848; AAC38254.1; "		
DR	HSSB: P13569; INBD.		
DR	InterPro: IPR003593; AAA.		
DR	InterPro: IPR001140; ABC_transporter_tmam.		
DR	InterPro: IPR003439; ABC_transport.		
DR	InterPro: IPR001687; ATP_GTP_A.		
DR	Pfam: PF00664; ABC_membrane. 1.		
DR	Pfam: PF00005; ABC_tran. 1.		
DR	SMART: SM00382; AAA. 1.		
DR	PROSITE: PS00211; ABC_TRANSPORTER. 1.		
DR	ATP-binding; Protease; Transport.		
QW	SEQUENCE 579 AA; 62531 MW; 54052E01DDE9E0D7 CRC64;		





RN [1]  
 RP NCB1\_TaxID=287;  
 RC  
 RC STRAIN=ATCC 15692 / PAOI;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wedman S., Yuan Y.,  
 RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAOI, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 CC -i SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
 CC (ABC TRANSPORTERS).  
 CC EMBL: AE004761; AAC06794.1; -.  
 DR HSSP, P13569; 1NEB.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR001140; ABC\_transporter\_tmam.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001687; ATP\_Grp\_A.  
 DR Pfam: PF00664; ABC\_membrane; 1.  
 DR Pfam: PF00005; ABC\_tran; 1.  
 DR SMART: SM00382; AAA; 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 KW ATP-binding; Complete proteome; Transport.  
 SQ SEQUENCE 600 AA; 63454 MW; D3DEA5FBA156FB2A CRC64;

Query Match	41.28;	Score 1156;	DB 16;	Length 600;
Best Local Similarity	43.48;	Pred. No. 2.3e-62;		
Matches 230;	Conservative 117;	Mismatches 181;	Indels 2;	Gaps 2.

Oy	20	KRAVITAMVESEFINILAVASVLYIMQVYDVKVTSKNSVLYVIVICVFLVIVISLEA	79
Oy	20	RAALRSVAFLEFTAVINILMLAPSLYIMQVYDVKVTSKNSVLYVIVICVFLVIVISLEA	79
Oy	80	LEPROVIVRGGLAFEDGVARPDIFKSVLSDTLSLR-KGIGAGAFRMDQVREPMNG-GLIAC	137
Oy	80	VMSLVIVIRGGQIDMOQJLONRITDASRASLERGEQAAGALNDVTSLRQPTIGNALFAFE	139
Oy	138	DAPWTFEVIYVSMKMLPEFGIILAIICIIIFGLAVNNDATKNPIMATWASIAONDAG	197
Oy	140	DAPWFLVLYLVIFLFSFPMJGLLALAGALLVLLAVNNSRSEPLAEAGLSITATQAS	199
Oy	198	STLRNAEVWKANGMGGLOARRRPARDEQVAMQAAASDAGANMSGIYFRNVOTLLIG	257
Oy	200	ANLROETTLAAMMLPAMRAMFQAQALFARONLTSERSAALGAASKGVRLAOSLVIG	259
Oy	258	GGAYVILIDKITSAGMIAGSILVGRALAPIEGAVGOMKNYIGARGAMDRLQYMLREKSA	317
Oy	260	LGAMVLAVDGLITGMIMIASILMGVRLSPIDOLJAVRWQMSGARQVGRRLRLLEENPPA	319
Oy	318	DDHMLPEPRGVLSAEASTLPPGAQOPTWRQASFRIDAGAVALVGPSAAGSSLLRGT	377
Oy	320	ALCPMPAPRAGALRVERLCAAPAREQALLODGFALPEBAGLVGPSGSGSTLARLL	379
Oy	378	VGVWPCAAGYIIRLDGVYDIKOMDEPKIGRHNGYVPODIETLPGSTVAONIAFTEESOEVI	437
Oy	380	VGAMQPLISGVARLDGADURQMSAALGPHTIGYLAQDVOLFAGSIAETIAEIAFAEYDAEKVY	439
Oy	438	EAATTLAGVHMIMOSLPMGVYDTAIGEGASLSSGOROLARLAVFERMPALLVDEPNASL	497
Oy	440	AAARLAGVHDLVLRPLPGYDTRLGDGAGLSGGORRIGLARALYGPALVILDEPNASL	499
Oy	498	DQYGEVALMAAMKRKAARTVIFATIKVULLQADYIYVINGOVISDFG	547
Oy	500	DEGEALALAIAMARQSGSLVLYTHKPAVLLATDCLLHLHGGRLORF	549

ID	054416	PRELIMINARY;	PRT:	582 AA.
AC	054416;			
DT	01-NOV-1996 (TReMBLrel. 01,	Created)		
DT	01-NOV-1996 (TReMBLrel. 01,	Last sequence update)		
DT	01-DEC-2001 (TReMBLrel. 19,	Last annotation update)		
DE	METALLOPROTEASE TRANSPORTER.			
GN	HASD8000.			
OS	Serratia marcescens.			
OC	Bacteria; Proteobacteria; gamma	subdivision; Enterobacteriaceae;		
OC	Serratia.			
OX	NCBI_Taxid=615;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SR41;			
RA	Omori K.;			
RT	"Cloning of the hasD8000 genes from <i>Serratia marcescens</i> Str41:			
RT	secretion mechanism for the lipase and metalloprotease ",			
RL	submitted (FEb-1996) to the EMBL/Genbank/DBJ databases.			
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.			
DR	EMBL; D83582; BAA12015.1; -.			
DR	HSSP; P13569; INBD.			
DR	InterPro; IPR003593; AAA.			
DR	InterPro; IPR001140; ABC_transporter_tmcm.			
DR	InterPro; IPR003439; ABC_transportr.			
DR	InterPro; IPR001687; ATP_GTP_A.			
DR	Pfam; PF00664; ABC_membrane; 1.			
DR	Pfam; PF00005; ABC_tran; 1.			
DR	SMART; SMO0382; AAA; 1.			
DR	PROSITE; PS00211; ABC_TRANSPORTER; 1.			
DR	ATP-binding; Protease; Transport.			
QO	SEQUENCE 582 AA; 62053 MW; 2FA93FEF300F3689F CRC64;			

Query Match	40.5%;	Score 1138;	DB 2;	Length 582;
Best Local Similarity	44.2%;	Pred. NO. 2.7e-61;		
Matches 232;	Conservative 105;	Mismatches 186;	Indels 2;	Gaps 2;

QY	28	VESFPIINLAIASPLYMLOQYDVRVTSNRVSLTIVLVCVLEFLVYGLLELKRIVLVR	8
Db	16	LEFTAVINILLMLPALPYLMQYDVRVLPSCGRMTLAMITLTVYGGLTMEGLLEWVRSSQVVR	75
QY	88	GGLRKEDGVARDPITFSVLDSTL-SRKRGIGGQAFRMDQVREGMT-QGLAFCDAPMTPE	145
Db	76	IGAOMDMRLNORVYDAEPFETNLKTGNPLNGALINDLTJNIRQFATGALAFDAPPEFY	135
QY	146	VIVSMMLHPFEFGILAIACIIIFGLAVMNDNTKNPIOMATASIAQNDAGSTLENAEV	205
Db	136	LLVLEFLHPLHWGLALASAGYIVLVLAMLNQVRSQAPLAEGRVALSATGOANGLNRAEA	195
QY	206	KKAMGMGGLQARWRARDEOYAWQAAASDAGAVNSGIKFVNIVOTLLIGGATLAD	265
Db	196	IAMGMGLTLRLKRWLNQHOOFLLQNRASEKIIAVYAWKVTYALQSLMLGGALLAVN	255
QY	266	GAISAGAMTAGSTLVGRALAPTEAGVQOMKNYIGARGAMDRLQTMREEKSADNHPDE	325
Db	256	GDTTPEBMTAGSTLIGRVLTGPIDQLIGANKQWSSANQSIQRLVEMLAANDPOAASLPLRA	315
QY	326	PRGVISAEEAASILIPGAQOPTRMQASFRIDAGAVALVCPESAAGKSSLRIGIVWPCAA	385
Db	316	PNGTLTVNOLASAPCGSAPVLHGVFSFRLEAGVELVIGASGGKTLMLQALGALTPLS	375
QY	386	GVYRLDGYOIKOMDEPKLRHVGVLRODLETFEGYVAONNARPTFEESOEVEIATLGV	445
Db	376	GDVRLDGLADIRQMDKOQLPGLYVLPDODIOLFGTLTENARRGVQDAELVVAAAALGV	435
QY	446	HEMIQSLPMGYDTAIEGGASLSGGGQRQRLARAARVFRMPALLVLDEPNASLDQVGEVAL	505
Db	436	HQIILHLPXYETELEGGSGLSGGQRQRYALARALAYGSPALVLLDEPNANLDREGEAL	495
QY	506	MEAMKRLKAARKVIFATHKVYNLLAQDYIMVINCQVISPFGGRD	550
Db	496	LRLIEALKRAGTIVIVTNRKPALLATDTKLLVLAAGOVQHFSGSD	540

RESULT 13  
Q54416

•  
•  
•  
•

RESULT 14  
ID 092NX3 PRELIMINARY; PRT; 617 AA.  
AC 092NX3;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE PUTATIVE TOXIN SECRETION ATP-BINDING ABC TRANSPORTER  
DE PROTEIN.  
GN SMC04207.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_Taxid=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21368234; PubMed=11474104;  
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,  
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,  
RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,  
RA Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,  
RA Groux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,  
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,  
RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,  
RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,  
RA Ramsberger U., Surzycki R., Thebaud P., Vandenbol M.,  
RA Vorhoefer F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;  
RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";  
RL Science 293:668-672(2001).  
DR EMBL: AL591789; CAC46614.1; -  
KM Complete proteome.  
SQ SEQUENCE 617 AA; 65748 MW; B7E513DEF98FD879 CRC64;

Query Match 38.8%; Score 1088.5; DB 16; Length 617;  
Best local Similarity 39.4%; Pred. No. 2.9e-58;  
Matches 220; Conservative 126; Mismatches 202; Indels 11; Gaps 6;  
QY 1 MFRRSGAKPIIFDOAVLV--ARRAVITAMVPSFFINITALVSPLYMQLVDRVITSRNV 58  
DB 1 MFGNSKRTTPOLKAAAVIPAKRTAFGTGAIISGIVNLALTSPLEMQVYDVYVLSRSLP 60  
QY TLVLYVYICFELFVLVGLLEALRTQVLVGRGLKFD---GVARPIKSVLDSLSKGI 114  
DB 61 TLVGLAVLATLMTGFMALIVSRALLRIGESFDHRLSGVHAATVRLPETHM--PGD 118  
QY 115 GGAQFRDMDOVREFMTG-GIARCDAPWTPFVIVSWMLHPFGILAIICIIIFGLAVM 173  
DB 119 GLDPLRDLNVGRFLAGAGFTALFDLPWMPLYLAICFLPHFWIGMPLVGLVAVVAMTLL 178  
QY 174 NDNAATKPIOMATYASTAANDAGSTLRNAEYKAMGMGGLARWRARDEOVAMQAAA 233  
DB 179 TDRLSVOPRTDARFYGTDRNALMEARRNAETVRALEGRLERHLOQRWLSAMHLYLEARRA 238  
QY 234 SDAGAAVMSGICFERNVOTLLIGGGAYLADKISAGAMIASIIVGRALPIEGAVGQ 293  
DB 239 ADAAGAGSISKRLRMLQSLILGAVMAYLEQOVNSAGVMAASSIMKGRALAPVDLAIGS 298  
QY 294 WKNYIGARGAMDRLOTMLREKSSADDMPLPERPGVLSAEASILTPGA--OOPTROASF 352  
DB 299 WKPEFLMARQSMDRLRNLDLIPRANRPMPLPAPKSIILAVESGLVYIPPGSGSGKPTVAQVSF 358  
QY 353 RIDAGAAVALVPSAAGSSILKGIQVWPCACVIRLDGDIKQMDPEKLGRNVGLPQ 412  
DB 359 SVPGGAGIIGIIPSGSGKSTTRALVGAMTPASGKVRLDGASYEQMDREELGRIIGYLPQ 418  
QY 413 DIELFSTVNONIATPEF--ESQEVIEATLLAGVHEMIQSLPMGYDAIIGGGSGGQ 471  
DB 419 GVELFDSTIAENISRFNSNDPTAIYAAQAGVHDLIVREQGYETRIIGAGSALSAGQ 478  
QY 472 RQRLATLAAVFRMPALLVLDEPNASLDQVEVALMEAMKRLKAKRTVIVATHRVNLAAQ 531

DB 479 RQRLIGARALYNDFLVLVDEPNANLDDAGGEKAVIDAITSVRNREGIAYVVAHRPSAIGA 538  
QY 532 ADYIMVINOQVISPGEKD 550  
DB 539 ADVLVVDSSGVNAFGPRD 557  
RESULT 15  
ID 033678 PRELIMINARY; PRT; 562 AA.  
AC 033678;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE PRSD (ABC TRANSPORTER ATP-BINDING PROTEIN).  
GN PRSD.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_Taxid=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA York G.M., Walker G.C.;  
RT "The Rhizobium meliloti exoK gene and prsd/prse/exsh genes encode components of independent degradative pathways which contribute to production of low-molecular-weight succinoglycans.";  
RL Mol. Microbiol. 25:117-134(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021; PLASMITD=PSYMB (MEGAPLASMID 2);  
RX MEDLINE=21396508; PubMed=11481431;  
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,  
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,  
RA Golding B., Puehler A.;  
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliloti.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).  
DR EMBL: U89163; AAB64091.1; -  
DR EMBL: AL603646; CAC49692.1; -  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR001140; ABC\_transporter\_tmam.  
DR InterPro: IPR003439; ABC\_transport.  
DR InterPro: IPR001687; ATP\_GTP\_A.  
DR Pfam: PF00664; ABC\_membrane\_1.  
DR Pfam: PF00005; ABC\_tran; 1.  
DR SMART: SM00382; AAA; 1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; UNKNOWN\_1.  
KM ATP-binding; Plasmid; Complete proteome.  
SQ SEQUENCE 562 AA; 62119 MW; 102FEF38C9ED6BA CRC64;  
Query Match 38.4%; Score 1077; DB 16; Length 562;  
Best local Similarity 42.2%; Pred. No. 1.3e-57;  
Matches 226; Conservative 104; Mismatches 202; Indels 4; Gaps 3;  
QY 20 RPAVITAMVPSFFINITALVSPLYMQLVDRVITSRNVSTIIVTVICVPLFYGLLEA 79  
DB 19 RAAITGVAVSALVNLILYLGSSFFMLEVYDRILPSRSIPSLIALSLALLAFQAGFEL 78  
QY 80 LRTQVLVRGGLKFPDGVARDPIFKSVLDSLS--RKGIGGAQFRDMDOVREFMTG-GIAR 136  
DB 79 IRGMALVRIAGALDESUNGRIYRAIVAPLKLKMGSGGLDALRFDQVRSLSGVGRAAL 138  
QY 137 CDAPWTPFVIVSWMLHPFGILAIICIIIFGLAVNNDNAATKPIOMATYASTAAND 196  
DB 139 FDLFWLPEFYLAICFLPHVIGLIIAIIIGLITLTLYLTNRQTPAPKASBAGGLRVFA 198  
QY 197 GSTLRNENVKAMGMGGLARWRARDEOVAMQAAASDAGAAVMSGIKVFNRNVQTLIL 256  
DB 199 QASQNRNEVVHAMGMSARLTAIWMERRNTERDENRRTSDINGYALSKYFRMALQSGVL 258  
QY 257 GGGAYLADKISAGAMIASIIVGRALAPIEGAVGQWKNYIGARGAMDRLOTMLREKSS 316

```

Db 259 AAGAVLVIRGEASPGIITIASILTARALAPVELAIGNMRGLVNAQSWORLKEELNALPE 318
QY 317 ADDHMLPEPRGVLSAEASSTLPPGAQOPTMRQASFRIDAGAVALVGPSSAGKSSILRG 376
Db 319 ADAPLOLPDHERLIVEGLASGPPAQRVSVDFNETVRAGAVGIVGPSAGKSSILARA 378
QY 377 IVGVWPCAGVIRLDGYDIKQWDPKLRHVGYLPQDIELFSGTVANQIARFTE-FESQE 435
Db 379 ILGIWPAVRSVRLDGDALDQWSDALGKHVGYLPQDVLEFAGTIAQNICRFADATSEA 438
QY 436 VIEAATLAGVHEMIQSLPMGYDTAIGBEGASLSCGQORLALARAVERFMPALLVDEPNA 495
Db 439 IVAAKAKAARYNDLILRLPNGDYDTEIGDGMTLISAGORVVALARALYGDPLVYDEPNS 498
QY 496 SLDOVEVALMEAMKRLKAKKRTVIFATHKVNLLAQADYIMVINOGVYISDGERDR 551
Db 499 NLDAEGQALSEAIMSVSRGIVYVAHRPSALASVDVLMNENGRMQAFGPREQ 554

```

Search completed: July 19, 2002, 10:48:41  
 Job time: 416 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:43:17 ; Search time 55.96 Seconds

(Without alignments)  
863.422 Million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165

Sequence: 1 MKPKIQRTDNFQAVARIG.....VLQYLESPLRDLRTTMRRE 435

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Database :

Listing first 45 summaries

```
1: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/genesqp-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2165	100.0	435	21	AA822787	Caulobacter cresce
2	521.5	24.1	443	17	AA805298	Esterase secretory
3	495	22.9	444	20	AA755920	Pseudomonas fluore
4	479.5	22.1	443	20	AA755923	Pseudomonas fluore
5	335	15.5	415	22	AB852753	Escherichia coli p
6	329	15.2	475	21	AA755573	Neisseria meningit
7	313.5	14.5	477	18	AAW22161	Apixiid protein.
8	313.5	14.5	477	21	AAW51414	A. pleuropneumonia
9	262	12.1	478	18	AAW22154	Apixid protein. Ac
10	262	12.1	478	21	AAW51408	A. pleuropneumonia
11	261.5	12.1	435	21	AAV75574	Neisseria meningit

12	201.5	9.3	425	20	AAV24069	Salmonella typhimu
13	149.5	6.9	543	22	AA667295	Amino acid sequenc
14	142	6.6	355	22	AA698387	Escherichia coli p
15	130	6.0	285	22	AAU36320	Pseudomonas aerugi
16	128.5	5.9	1761	20	AAV15457	Human laminin beta
17	126.5	5.8	412	21	AAV90283	N. meningitidis BA
18	125.5	5.8	1839	21	AAV94291	Caenorhabditis ele
19	124.5	5.8	659	22	AA693061	C glutamicum prote
20	122.5	5.7	2779	22	AB862371	Drosophila melanog
21	119	5.5	331	20	AAV34473	Porphyromonas ging
22	119	5.5	365	20	AAV34349	Porphyromonas ging
23	118.5	5.5	796	20	AAW32798	S. aureus SecA2 pr
24	117.5	5.4	233	22	AAW3878	Human polypeptide,
25	116	5.4	1154	22	AAU32042	Novel human secret
26	115.5	5.3	1294	22	AB863502	Drosophila melanog
27	115.5	5.3	2472	22	AAU33568	Pseudomonas aerugi
28	115	5.3	1090	21	AAV99270	Human huntingtin-1
29	115	5.3	2192	18	AAW21732	LexA/NuMA fusion p
30	115	5.3	2272	18	AAW21731	GAL4/HA/NuMA fusio
31	114.5	5.3	390	22	AAU35553	Haemophilus influe
32	114.5	5.3	1177	22	AA896721	Putative P. abyssal
33	114.5	5.3	1717	22	AB820672	Novel human diagno
34	114	5.3	1374	22	AA869070	Human male enhance
35	113.5	5.2	399	21	AAV81743	Streptococcus pneu
36	113.5	5.2	457	20	AAW94867	Accessory protein
37	113.5	5.2	541	22	AAW94625	Human protein sequ
38	113.5	5.2	1201	22	ABG10512	Novel human diagno
39	113.5	5.2	1235	22	ABG24756	Novel human diagno
40	113.5	5.2	1283	22	ABG14964	Novel human diagno
41	113.5	5.2	2117	22	AAU32040	Novel human secret
42	113	5.2	2400	22	ABG20278	Novel human diagno
43	113	5.2	2415	22	ABG20279	Novel human diagno
44	113	5.2	792	22	AAU34228	Staphylococcus aur
45	112.5	5.2				

#### ALIGNMENTS

```
RESULT 1
AAB22787 standard; protein: 435 AA.
ID AAB22787:
AC AAB22787:
XX
DT 22-DEC-2000 (first entry)
XX
DE Caulobacter crescentus membrane fusion protein (MFP), RseE.
XX
KW Membrane fusion protein; MFP; RseE; channel protein;
KW type I secretion system; C-terminal secretion signal;
KW Caulobacter host cell; biofilm formation; bioeffector;
KW recombinant protein expression host; sewage treatment;
KW wood pulping; fish vaccine.
XX
OS Caulobacter crescentus.
XX
PN W0200049163-A1.
XX
PD 24-AUG-2000.
XX
PF 21-FEB-2000; 2000WO-CA00173.
XX
PR 19-FEB-1999; 99CA-2261186.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
XX
PI SmIt J;
XX
DR WPI: 2000-571987/53.
XX
PT Caulobacter host cell for expression and secretion of a heterologous
polypeptide, useful for treating sewage, waste water and in the pulping
```

PT process -  
XX  
PS Claim 1: Page 35; 46pp; English.

CC The invention relates to a Caulobacter host cell for expression and  
CC secretion of a heterologous polypeptide. The Caulobacter host cell  
CC comprises at least one surface layer transport protein having an  
CC amino acid sequence homologous to the Caulobacter crescentus Rsad or  
CC RsaA proteins (AA822786, AA822787). The host cell further comprises a  
CC DNA construct comprising DNA encoding a polypeptide heterologous to a  
CC surface layer (S-layer) protein of the cell, from and operably linked  
CC to, a DNA encoding a Caulobacter surface layer protein secretion signal,  
CC with the proviso that when the cell comprises transport proteins having  
CC the same sequence as both the Rsad and RsaA proteins, the secretion  
CC signal is not from the Caulobacter crescentus surface layer protein  
CC (RsaA). The invention also encompasses a method for identifying a  
CC Caulobacter suitable for use as a host cell for expression and secretion  
CC of a heterologous polypeptide comprising the detection of homologues of  
CC the rsad or rsaa genes using a rsad/rsaa hybridisation probe. The host  
CC cell is used for the expression and secretion of a heterologous  
CC polypeptide. The modified Caulobacter cells may be used to treat sewage  
CC and waste water. The cells may also be grown in wood pulp suspensions and  
CC then used in the wood pulping process. They can also be used in fish  
CC vaccines. Caulobacter species are able to form biofilms, and can attach  
CC themselves to surfaces without producing the extracellular enzymes or  
CC polysaccharide slimes that are characteristic of most other  
CC surface-attached bacteria. Caulobacter are therefore particularly suited  
CC for use in bioreactor systems. The present sequence represents the  
CC Caulobacter crescentus membrane fusion protein (MFP). RsaA. This protein  
CC is anchored in the inner membrane of Caulobacter crescentus and is  
CC thought to span the periplasm. Together with an outer membrane protein,  
CC it forms a channel that extends from the cytoplasm through the two  
CC membranes to the outside of the cell.

XX  
SQ Sequence 435 AA;

Query Match 100.0%; Score 2165; DB 21; Length 435;  
Best Local Similarity 100.0%; Pred. No. 2,6e-178;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPPIQPTPTNFQAVNATIGIITLVGGLGNMAFPADDSAVANGVSAEVSQDQH 60  
DB 1 mkpkiqptptnfqavnatigiiitlvvgllgnmafpadsavlangvsevsqdvqh 60  
QY 61 LEGGLAKILVREGSEKAVAGVLELDPTQANAAGITRNCYVALKAMEARLAEEDQRP 120  
DB 61 legglakilvregsekvagvyleldptqanaaagitrncyvalkamearliaeerdqr 120  
QY 121 SISFPADLTQSADPMVARAIADQAOFTERRQTIGQVDLMNAQRLQYQSEIEGIDRQT 180  
DB 121 sisfpadltqsadpmvaraiadegaqfterrtqitgqvdlmnaqrlyqyseiegidrt 180  
QY 181 OGLKQULFIEDLIDLRKLDKLVGPPRRLALPARAGSLSGRLTADSKAVQAS 240  
DB 181 qglkqulfigedlidrlklydgviprrllalaeargsisgsigriladskavqas 240  
QY 241 DTQLVRIKIOFFQVOSISIFETRRLAEVTEKEVNASDAQRIKIVSPVNGTQONLRF 300  
DB 241 dtqlvrikiqffqvossisifetrriaevtekevasdaqrlikivspvngtqnlrf 300  
QY 301 FTGAVNRAAEPLVDIADDEAFVIOAHFOPTDVNVNIMGMWTEVRLPAFHSAGNDPDR 360  
DB 301 ftegvnraaeplvdiadpedafviahfopdvndvnmgmwtevrllpafhsagndpdr 360  
QY 361 HDPAVAVADRISSPOKQARFLGLIVRDVAKQLPRLRGRTAGMPAQVYTPTEGRTYQL 420  
DB 361 hdpavavadrispokrqlrflglivrdvdkqlprrlrgrtagmpaqvypvtegrtyql 420  
QY 421 FSPRLDTLRTTMRRE 435  
DB 421 fsprltdltlrmree 435

RESULT 2

AAW05298  
ID AAW05298 standard; Protein; 443 AA.

XX  
AC AAW05298;

DT 30-DEC-1996 (first entry)

XX  
DE Esterase secretory protein (ORF2 product).

XX  
KW Esterase secretory gene; protein secretion.

OS Serratia marcescens strain S141 (FERM BP-487).

XX  
PN EP733707-A2.

XX  
PD 25-SEP-1996.

XX  
PE 20-MAR-1996; 96EP-0104429.

XX  
PR 23-MAR-1995; 95JP-0063772.

PA (TANA ) TANABE SEIYAKU CO.

XX  
PI Akatsuka H, Kawai E, Shibatani T;

XX  
DR WPI; 1996-444599/45.

XX  
DR N-PSDB; AAT39662.

PT New isolated esterase secretory gene from Serratia - used to  
PT increase the productivity and extracellular secretion of esterases

XX  
PS Claim 3; Page 21-22; 29pp; English.

CC 3 Different proteins (AAW05297-99) participate in the mechanism of  
CC secretion of esterase by Serratia marcescens S141. They are  
CC encoded by an esterase secretory gene (AAT39662) that includes 3  
CC open reading frames. The proteins can be produced by transformed  
CC host cells, pref. S. marcescens or E. coli, carrying vector  
CC plasmids incorporating the esterase secretory gene. The esterase  
CC is secreted from the host cell, and is recovered from the culture  
CC medium and from within the cells. The esterase is useful in  
CC hydrolysis reactions.

XX  
SQ Sequence 443 AA;

Query Match 24.1%; Score 521.5; DB 17; Length 443;  
Best Local Similarity 29.3%; Pred. No. 1.5e-36;  
Matches 133; Conservative 88; Mismatches 194; Indels 39; Gaps 7;

QY 6 IORPDNF-----QAVARIGYGIITLVGGLGNMAFPADDSAVANGVSAEVS-Q 56  
DB 5 ipepdnsyteeipgderfrftmgwlvvgilgflawaafapldkvaspsvtsvgnrk 64  
QY 57 DVQHLEGMLAKILVREGSEKAVAGVLELDPTQANAAGITRNCYVALKAMEARLAE 116  
DB 57 dvqhlegmlakilvregsekvagvyleldptqanaaagitrncyvalkamearl 116  
QY 117 DQRPISFPADLTQSADPMVARAIADQAOFTERRQTIGQVDLMNAQRLQYQSEIEG 176  
DB 117 dqrpisfpadltqsadpmvaraiadegaqfterrtqitgqvdlmnaqrlyqyseieg 176  
QY 176 gqisvltisplidavkdprvaeilaqlqlfastergaidsid-----gykxsmgdi 177  
DB 176 gqisvltisplidavkdprvaeilaqlqlfastergaidsid-----gykxsmgdi 177  
QY 177 DRQTQGLKD-----QLGFIEDLIDLRKLDKLVGPPRRLALPARAGSLSGRLT 229  
DB 178 rftqlvgldsrnqkqldglsiregmnmkqlaadylpnrylervqgrfaevnsidelv 237  
QY 230 ADRAKAVGASDPTQKVRQIKQEFFQVOSISIFETRRLAEVTEKEVNASDAQRIKIVS 289  
DB 238 grtqgkqgldesqgrldqrfadygrevrftqagcmadasefrnkqgmadfgnltats 297  
QY 290 PVNGTRQONLRFTEGAVNRAAEPLVDIADDEAFVIOAHFOPTDVNVNIMGMWTEVRLPA 349



Db 298 pvgctvgvgnlftcgvgvvgagdhlmvdpqcalivvdsrlkvglfdkvygplvdimftca 357  
QY 350 FHAGAPDEPERHDPVAVADRIDSP-----PQKQARLFL---GIYRVQVKQPLRLRGKRVTA 401  
Db 358 fgnqtkpklpvtclvsadrlvdkangepyygmqvsvpegmkmlsgedlkr----- 409  
QY 402 GMPAQVIVPTGERTVLQYLFSPRLDRLRTTMRRE 435  
Db 410 gmpvevfvtktgrslslsylvlfpklldrahcsltec 443

## RESULT 3

AA55920  
ID AAY55920 standard; Protein; 444 AA.

AC AAY55920;

DT 15-FEB-2000 (first entry)

DE Pseudomonas fluorescens ABC transporter cassette component B.

KM ATP-binding cassette; transporter; operon; lipBCD; Serratia marcescens;

KW microbe; protein secretion.

OS Pseudomonas fluorescens.

PN JP11276172-A.

PD 12-OCT-1999.

PE 27-MAR-1998; 98JP-0080597.

PR 27-MAR-1998; 98JP-0080597.

PA (TANA ) TANABE SEIYAKU CO.

DR WPI: 1999-626936/54.

DR N-PSDB: AA622700.

PT An ABC transporter gene - used to increase the ability of a microbe to secrete proteins

PS Claim 2; Page 16-17; 28pp; Japanese.

CC This sequence represents an ATP-binding cassette (ABC) transporter component B from an Pseudomonas fluorescens strain 33 and encoded by an ABC transporter operon. The operon comprises 3 genes where the termination codon of the first gene overlaps with the initiation codon of the second gene. The sequences of the encoded proteins have 60, 44 and 46% homology respectively to the lipBCD proteins from Serratia marcescens. The novel gene and protein can give or increase the ability of a microbe for secreting a protein.

XX Sequence 444 AA;

Query Match 22.9%; Score 495; DB 20; Length 444;

Best Local Similarity 30.9%; Pred. No. 2.9e-34;

Matches 134; Conservative 86; Mismatches 201; Indels 12; Gaps 7;

QY 9 PTFNFOAVNRIGIGIALTFVGLLG---WNAFAPLDSAVIANG-VVSAEVSQDVQHLGG 64

Db 18 pelgaraffarmgv---lllvvgaggfllwaslapldqglpvggtlvvsgkrkavqlspg 74

QY 65 MLAKILVRGEGEKVKAQVLFELDPTQANAAAGITRNQVYALKAMEARLLAERDQPSISF 124

Db 75 vvsrllvreegaevkqgplfrldqtngqdvshlgqymawasvawqserdngstlft 134

QY 125 PADUTSQRADPMYARALADEQAQOFTERROTIGQVULMNAORLQYSEISGIDROTGLK 184

Db 135 pel-sgnppqatalvleqgrqlfssrteafareqagltanlegataqlngmrarsdlit 193

QY 185 DLQGFIEDELIDRLKLYDKGLVPRRLALEARAGSLSGSIGHRLTADRSKANQASDTOL 244  
Db 194 aqagslrdqlnmlqpladngyliprnrlmeyqrlsqvqqlaenfgsgrveggllsle 253  
QY 245 KVRQIKOEFEFEYQSITTEFRVRLAEVTEKEVVASDAQRIKIVSPVNGTAONLRFPTBG 304  
Db 254 klqhsheeykvevrsqldaqqlrsltleqqltsagfdlqhselnadpqlavnlsvhteg 313  
QY 305 AAVRAAEPLVIDIAPDEEAFVIOAHFOPTVDVNVHMGVTFEVRILPAFHSGNPREHDPV 364  
Db 314 avveragecllelvpqdrlevegrlpvhlvdkvghlhpvdlftalnqstrprvgsal 373  
QY 365 AVADRISDPQKQARLFLGIYRVQVKOLP--PHLRGV-TAGMPAQVIVPTGERTVLQYLS 422  
Db 374 lsadgmldektgmpy--vlttvssasaleklnlgvlkpgmpemefirgersllylfk 431  
QY 423 PLRDTLRTTMRRE 435  
Db 432 plldragsaltec 444

## RESULT 4

AA55923  
ID AAY55923 standard; Protein; 443 AA.

AC AAY55923;

DT 15-FEB-2000 (first entry)

DE Pseudomonas fluorescens ABC transporter cassette II component B.

KM ATP-binding cassette; transporter; operon; lipBCD; Serratia marcescens;

KW microbe; protein secretion.

OS Pseudomonas fluorescens.

PN JP11276172-A.

PD 12-OCT-1999.

PE 27-MAR-1998; 98JP-0080597.

PR 27-MAR-1998; 98JP-0080597.

PA (TANA ) TANABE SEIYAKU CO.

DR WPI: 1999-626936/54.

DR N-PSDB: AA622701.

PT An ABC transporter gene - used to increase the ability of a microbe to secrete proteins

PS Claim 8; Page 23-24; 28pp; Japanese.

CC This sequence represents an ATP-binding cassette (ABC) transporter component B from an Pseudomonas fluorescens strain 33 and encoded by the ABC transporter operon II. The operon comprises 3 genes where the termination codon of the first gene overlaps with the initiation codon of the second gene. The sequences of the encoded proteins have 63, 56 and 59% homology respectively to the lipBCD proteins from Serratia marcescens. The novel gene and protein can give or increase the ability of a microbe for secreting a protein.

XX Sequence 443 AA;

Query Match 22.1%; Score 479.5; DB 20; Length 443;

Best Local Similarity 29.0%; Pred. No. 6.3e-33;

Matches 130; Conservative 83; Mismatches 196; Indels 39; Gaps 7;

QY 9 PTFNFOAV-----ARIGYGIALTFVGLGMAAFAPLDSAVIANG-VVSAEVSQDVQHL 62

Db 14 pvnvialddkkyarlgwlllvlgfagflgwaalaplkdgvavsgkvmvshktvqhps 73

[illegible]

```

5 RESULT 5
ABBS2753 ID ABB52753 standard; Protein: 415 AA.
XX
AC ABB52753;
XX
DT 11-FEB-2002 (first entry)
XX
DE Escherichia coli polypeptide SEQ ID NO 907.
XX
KM Escherichia coli; B2/D+A-; antinflammatory; antibacterial;
KM immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
KM systemic infection; non-diarrhoeal infection; septicaemia;
KM pylonephritis; antibiotic resistance.
XX
OS Escherichia coli.
XX
PN W0200166572-A2.
XX
PD 13-SEP-2001.
XX
PF 12-MAR-2001; 2001WO-EP03445.
XX
PR 10-MAR-2000; 2000FR-0003145.
XX
PR 02-FEB-2001; 2001FR-0001449.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Bingen E, Bonacorsi S, Clermont O, Nassif X, Tinsley C;
DR WPI; 2001-550253/61.
XX
PT A library of DNA fragments of Escherichia coli strains for the
PT phylogenetic determination of a given strain comprises polynucleotides of
PT nature B2/D+ A-
XX
XX Example 6; Fig 6; 646bp; English.
XX
PS
XX
CC The invention relates to a library of DNA fragments of Escherichia coli
CC strains comprising polynucleotides (ABA88577-ABAB8729 and ABAB9533)
CC and encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature
CC B2/D+A-. The polynucleotides have potential antinflammatory,
CC antibacterial and immunosuppressive activity as part of pharmaceutical

```

CC compositions used to treat, palliate or prevent extra-intestinal *E. coli*  
CC infections. The polypeptides are useful for determining the phylogenetic  
CC group of a given *E. coli* strain. These polypeptides can detect and treat  
CC an undesired development of *E. coli*, particularly an extra-intestinal  
CC infection that include systemic and non-diarrhoeal infections such as  
CC septicaemia, pyelonephritis and meningitis this is particularly  
CC advantageous as bacterial resistance is increasing with the more  
CC frequent use of broad spectrum antibiotics.

Query Match	15.5%	Score 335	DB 22	Length 415
Best Local Similarity	25.6%	Pred. No. 1.7e-20		
Matches 110	Conservative 74	Mismatches 162	Indels 84	Gaps 11
23	IIATLFFVGLGMAAPAPLDASAVIANG-VVSAEVSODVQHLLEGGMALAKILVREGGEKKAGQ	81		
Db	47 VLVTVFV---LkwmSpIdevtrggstIpgsregyIqIdIpqIILktleIvreggdIvekgq	103		
Qy	82 VFELDPTQANNAAGITRNQVYALKAMEARLLAERDQPSISFPADLTISORADPMVVARAI	141		
Db	104 VLLTIDTTSsamlresearvnlnleavrarIraea-ysestIfpddv-----	149		
Qy	142 ADEQAFTRRRRTTIOGVNDLMAQRLOQYSELEGIDRQCGKLDQGFTEDELIDLRKIX	201		
Db	150 ---pedIteRestv-----yLrktleIagsIagIkyskaIlIdkeIamtlrIv	193		
Qy	202 DGIAPRPRLRLLEARAGSLSGISGRLTGRSKAVOGASDPTOLKVRQIKOEFEFGYSQSI	261		
Db	194 rgaImseveellImqgsaeI-----qIgmdekpnkylteegael	232		
Qy	262 TETRVRLAEVTEKEVVASDAQRIKIVSPVNGTAQNLRFTEGAVRAAEPLVDIAPEDE	321		
Db	233 VkteeelaqakemagradpverstrIrapIrgIvknItrvnlIgyvsaqgdImeiIplred	292		
Qy	322 AVVIAQAPRPDTPVDNVHMKMVTENVLRPAHSA-----GNDRPRRHDPVANADRI	370		
Db	293 qIlIleayIprIdvayvrtgmpaIvktIdaydIaygIdgVvnlvsbdlr-dqkIrggdIk	351		
Qy	371 SPPOKARLFLGIVRVYDKQLRPHLRGR-----VYAGMPAGVYPTGERTVLAQYLFSP	424		
Db	352 Idpne-----ayvIvltlsmnYltdrngkIlprvIpgmaIasvdlktgqksvfygIkpri	405		
Qy	425 RDLRTTMRK 434			
Db	406 -----trmkq 410			
RESULT 6				
AAV75573				
ID . AAV75573 standard; Protein; 475 AA.				
AC AAV75573;				
DT 21-MAR-2000 (first entry)				
DE Neisseria meningitidis ORF 764 protein sequence SEQ ID NO:2620.				
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;				
KM antigenic; diagnosis; immunogenic; Infection; meningitis; septicemia;				
KX antibacterial; gene therapy.				
OS Neisseria meningitidis.				
PN WO957280-A2.				
XX 11-NOV-1999.				
XX 30-APR-1999; 99WO-US09346.				
XX 01-MAY-1998; 98US-0083758.				
XX 31-JUL-1996; 98US-0094869.				

PR 02-SEP-1998; 9805-0098994.  
PR 02-SEP-1998; 9805-0099062.  
PR 09-OCT-1998; 9805-0103749.  
PR 09-OCT-1998; 9805-0103794.  
PR 09-OCT-1998; 9805-0103796.  
PR 25-FEB-1999; 9905-0121528.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (GENO-) INST GENOMIC RES.  
XX  
PI Eraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,  
PI Petersen J, Pizzo M, Rappelli R, Ratti G, Scalato E, Scarselli M,  
PI Tettein H, Venter JC;  
XX  
DR MPI: 2000-062150/05.  
DR N-PSDB; AA254335.  
XX  
PT Novel Neisserial polypeptides predicted to be useful antigens for  
PT vaccines and diagnostics -  
XX  
PS Claim 2; Page 1243; 1453pp; English.  
XX  
CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941  
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides  
CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent  
CC PCR primers used in the exemplification of the present invention. The  
CC polypeptides, the polynucleotides, antibodies and compositions of  
CC the invention can be used as vaccines, as diagnostic reagents, and as  
CC immunogenic compositions. The polypeptides can be used in the  
CC manufacture of medicaments for treating or preventing infection due to  
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the  
CC presence of Neisseria bacteria, or to raise antibodies. They may also  
CC be used to screen for agonists or antagonists, which may themselves  
CC have use as antibacterial agents. The polynucleotides of the invention  
CC may also be used in gene therapy protocols.  
XX  
SQ Sequence 475 AA:

Query Match 15.2%; Score 329; DB 21; Length 475;  
Best Local Similarity 26.2%; Pred. No. 6.7e-20;  
Matches 128; Conservative 78; Mismatches 188; Indels 94; Gaps 18;  
OY 1 MKPPKIORPTDNF-----QAVARIQYGIILTFVGLGMAFAPLDSAV 44  
DB 27 lhpkttaeagafllphlcltdtprsaapkaarf---lmafallll wswfgkldiya 82  
OY 45 IANG-VVSAEVSQDVQHLLEGMLAKILVREGKVKAGVLFELDPT-----QANA 93  
DB 83 aasgktvsggrsktlqplletavkvahvrdgqhvkggetlaeavgtldsvvsgalq 142  
OY 94 AAGITNNOYVA-LKAMEARLAEERDQRPSTSPFADITSGRADPMVARALDEAOQTERR 152  
DB 143 aaglsklyeavlaalestr-----tphldmagarslqslsdavqsag 185  
OY 153 QTIQGVDMNNAORLYOSEIEGIDRTOGLKQD-----IGFIEFE-LIDLRKLVDKGL 205  
DB 186 vlahqhygwaagdaqlgqalryhgaelsakageklysvgaieqgkctadyrrlladnf 245  
OY 206 VPRPRLLEARAGS---LSSIGRLTADRSKAVAGASDTQIKVROIKOEFE---QVS 258  
DB 246 lsehafllegqsksvnmwldestrgmrqlgaalaqaenrvintnklrtdlidalrqn 305  
OY 259 QSTTERRVLAETKEEVVASDAQRIKIVSPYNGTAQNLRFTEGAVVRAAPLVADIAP 318  
DB 306 eqldqyrgq---tdk---akgrqqlmtlqspadglvqelatltygvvgaadkmmvlap 358  
OY 319 EDEAFVIAQHFOTPDVNDVNHMGWTEVRLPAF-----HSAGPDEPERHDPVAVDRISD 372  
DB 359 dddkmdvevrlvlnkdigfegqgdavvkklesfpyltygyltgkvvksndavs-----411  
OY 373 POKQARLFLGIVRVNDVKOLPRH---LRGR---VTAGMPAOVIVPTGERTVLOYLFSPLRD 426  
SQ Sequence 475 AA:

DB 412 -heg----lglytavwsldkhtlndgkavnltaqnmvtaelktgkrvldy1llep1q6 466  
OY 427 TLRTTMR 434  
DB 467 kidesire 474  
RESULT 7  
AAW22161  
ID AAW22161 standard; Protein; 477 AA.  
XX  
AC AAW22161;  
XX  
DT 16-FEB-1998 (first entry)  
XX  
DE ApxIIID protein.  
XX  
KW RTX toxin; apxICA gene; apxIIB gene; apxIIABCD gene;  
KW repeat in toxins toxin; cell-associated RTX toxin; vaccine production;  
KW therapy; A. pleuropneumoniae infection; swine pleuropneumonia.  
XX  
OS Actinobacillus pleuropneumoniae.  
XX  
PN CA2170839-A.  
XX  
PD 02-SEP-1996.  
XX  
PE 01-MAR-1996; 96CA-2170839.  
XX  
PR 01-MAR-1995; 9505-0396244.  
XX  
PA (UYGU-) UNIV GUELPH.  
XX  
PI MacInnes J, Mallard B, Ricciatti P, Rosendal S;  
XX  
DR MPI: 1997-245536/23.  
DR N-PSDB; AAT73220.  
XX  
PT Preparations of microorganisms producing cell-associated RTX toxins  
PT especially for production of vaccines against swine  
PT pleuro-pneumonia  
XX  
PS Disclosure; Pages 113-114; 151pp; English.  
XX  
CC AAW22151-W22161 represent A. pleuropneumoniae RTX (repeat in toxins)  
CC toxins. These sequence are encoded by the apxICA, apxIIB, apxIIABCD, and  
CC apxIIABCD genes (see AAT73217-T73220), and can be expressed by  
CC microorganisms used in the preparations of the invention. The  
CC preparations are bacterial preparations comprising one or more isolated  
CC and purified strains of a microorganism that produces one or more RTX  
CC toxins, where the strains have at least one cell-associated RTX toxin.  
CC The preparations are used for production of vaccines for the prophylaxis  
CC and treatment of infectious diseases caused by microorganisms that  
CC produce RTX toxins, where the strains have been attenuated or  
CC inactivated. The vaccines are preferably against Actinobacillus  
CC pleuropneumoniae infection (swine pleuropneumonia). It has been found  
CC that A. pleuropneumoniae produces significant quantities of  
CC cell-associated RTX toxins when cultured under certain conditions, and  
CC that the whole-cell protein composition of the cultures corresponds to  
CC the whole-cell protein profiles obtained from cells recovered at  
CC necropsy from the pleural fluid of infected swine. Vaccination with a  
CC bacterin prepared from heat-inactivated cultures having significant  
CC quantities of cell-associated RTX toxins give significant protection of  
CC swine against challenge with homologous strains.  
XX  
SQ Sequence 477 AA:

Query Match 14.5%; Score 313.5; DB 18; Length 477;  
Best Local Similarity 25.3%; Pred. No. 1.5e-18;  
Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13;  
OY 19 IGYGIATLTFVGLGMAFAPLDSAVIANG-VVSAEVSQDVQHLLEGMLAKILVREGKVK 77

```

Db      60 laylimifflaiy-1s1skveivaasatgklvfsgnskelpienalvxdlvykdgqfv 118
QY      78 KAGOVLEL-----DPTQANNAAGITRNQYVALKAMEARLAEERQPSISF-PADLFT 129
Db      119 ekqglllnltaigcdadkqktkslglerldgyrkslysi--ehnrplldngadfd 176
QY      130 S-GRADPMVARAIADE-----QAFTERROTIGOVDMNAORLOYSEIEGIDR 178
Db      177 svgeedktgarrhlitegfewqkqykelayqrkgaekqvtlanlrkysasrie---- 232
QY      179 QTOGLKQQLGFIEDLDLKLIVKGLVPRRLALAEAR---AGSISGSGRTADRSK 234
Db      233 -----keklsdtkklygvksiskhellagenryveasneelsvysghkveesd 280
QY      235 AVOGASPTOLKROIKOEFPVOSQSTERRPRLAEVTEKEVNASDQKRK--IVSPVN 292
Db      281 lkqgedklvtqtkfksdllekigqnlgreqkltlelke-----qrqlasllraps 334
QY      293 GTAQNLRFTEGAVVRAAEPLVDIAPDEDAFVIOAHFQPTDVNVHMGWTEVRLPAFHS 352
Db      335 gtvgqlkthekgvttaetlmvlapeddvlevalignkdvgfvelgqeevklvetfpy 394
QY      353 AGNDPDERHDPVAVADRISDPQKARLFLGIVRVQKQLPRHLRG-RVTAGMPAQVIVPT 411
Db      395 lrygylgkvkltldalehpq-Iglvfnslieinkkltldgdkelqsgmsviaelkt 453
QY      412 GERTVLOYLFSPURDLRTTMRE 434
Db      454 gersvisfllspleesitesire 476

```

```

RESULT 8
AAV51414
ID      AAV51414 standard; protein; 477 AA.
XX
AC      AAV51414;
XX
DT      05-MAY-2000 (first entry)
XX
DE      A. pleuropneumoniae apxIIID protein.
XX
KW      RTX toxin; vaccine; Repeats in Toxins; immunostimulatory; antimicrobial;
        antiinflammatory; antiaerthetic; antiabortive; treatment; pneumonia;
        pleuropneumonia; septicemia; nephritis; arthritis; endocarditis;
        shipping fever; abortion; whooping cough; sleepy foal disease; joint ill;
        urinary infection; peritonitis; meningitis; gastroenteritis.
KM      passive immunization; apxIIID.
XX
OS      Actinobacillus pleuropneumoniae.
XX
PN      US6019984-A.
XX
PD      01-FEB-2000.
XX
PF      23-DEC-1996; 96US-0772270.
XX
PR      01-MAR-1995; 95US-0396244.
XX
PA      (UYGU-) UNIV GUELPH.
XX
PI      Mallard B, Rosendal S, MacInnes J, Ricciatelli P;
XX
DR      WPI: 2000-146864/13.
XX
DR      N-PSDB; AA288587.
XX
PT      Bacterial preparation comprising microorganisms which produce a member
        of the Repeats in Toxins (RTX) family, useful for treating swine
        pleuropneumonia, arthritis in swine, shipping fever and abortion in
        cattle, and sleepy foal disease -
XX
PS      Disclosure; Column 89-98; 96pp; English.
XX

```

```

CC      CC This invention describes a novel bacterial preparation (I) which
CC      CC comprises one or more isolated and purified strains(s) of a microorganism,
CC      CC cultured in tryptone yeast extract (TYE) broth, which produces one or
CC      CC more RTX toxins (belonging to the family of toxins referred to as Repeats
CC      CC in Toxins), where the strain(s) have at least one RTX toxin which is
CC      CC cell-associated. The products of the invention have immunostimulatory,
CC      CC antimicrobial, antiinflammatory, antiaerthetic and antiabortive activity.
CC      CC The bacterial preparation may be used as vaccines for the prophylaxis and
CC      CC treatment of infectious diseases caused by strains of microorganisms
CC      CC which produce one or more RTX toxins. The infectious diseases are swine
CC      CC pleuropneumonia, pneumonia, septicemia, nephritis and arthritis in swine;
CC      CC septicemia, nephritis, endocarditis and arthritis in piglets; shipping
CC      CC fever and abortion in cattle; whooping cough, sleepy foal disease or
CC      CC joint ill (purulent nephritis, arthritis) in foals; septicemia,
CC      CC polyarthritis and abortion in horses; and urinary infections,
CC      CC peritonitis, meningitis, and gastroenteritis. The bacterial preparations
CC      CC may also be used to prepare antibodies which may be used as a means of
CC      CC passive immunization. This sequence represents the Actinobacillus
CC      CC pleuropneumoniae apxIIID protein described in the method of the
CC      CC invention.
SO      Sequence 477 AA;

```

```

Query Match 14.5%; Score 313.5; DB 21; Length 477;
Best Local Similarity 25.3%; Pred. No. 1.5e-18;
Matches 112; Conservative 88; Mismatches 190; Indels 53; Caps 13;

```

```

QY      19 IGYGIALTVEGLGMAFAPLDSAVANG-VVSAEVSQDVQHHEGMLAKILVREGKV 77
Db      60 laylimifflaiy-1s1skveivaasatgklvfsgnskelpienalvxdlvykdgqfv 118
QY      78 KAGOVLEL-----DPTQANNAAGITRNQYVALKAMEARLAEERQPSISF-PADLFT 129
Db      119 ekqglllnltaigcdadkqktkslglerldgyrkslysi--ehnrplldngadfd 176
QY      130 S-GRADPMVARAIADE-----QAFTERROTIGOVDMNAORLOYSEIEGIDR 178
Db      177 svgeedktgarrhlitegfewqkqykelayqrkgaekqvtlanlrkysasrie---- 232
QY      179 QTOGLKQQLGFIEDLDLKLIVKGLVPRRLALAEAR---AGSISGSGRTADRSK 234
Db      233 -----keklsdtkklygvksiskhellagenryveasneelsvysghkveesd 280
QY      235 AVOGASPTOLKROIKOEFPVOSQSTERRPRLAEVTEKEVNASDQKRK--IVSPVN 292
Db      281 lkqgedklvtqtkfksdllekigqnlgreqkltlelke-----qrqlasllraps 334
QY      293 GTAQNLRFTEGAVVRAAEPLVDIAPDEDAFVIOAHFQPTDVNVHMGWTEVRLPAFHS 352
Db      335 gtvgqlkthekgvttaetlmvlapeddvlevalignkdvgfvelgqeevklvetfpy 394
QY      353 AGNDPDERHDPVAVADRISDPQKARLFLGIVRVQKQLPRHLRG-RVTAGMPAQVIVPT 411
Db      395 lrygylgkvkltldalehpq-Iglvfnslieinkkltldgdkelqsgmsviaelkt 453
QY      412 GERTVLOYLFSPURDLRTTMRE 434
Db      454 gersvisfllspleesitesire 476

```

```

RESULT 9
AAW22154
ID      AAW22154 standard; Protein; 478 AA.
XX
AC      AAW22154;
XX
DT      16-FEB-1998 (first entry)
XX
DE      ApxID protein.
XX
KW      RTX toxin; apxICA gene; apxIIBD gene; apxIIB/C gene; apxIIABCD gene;
        repeat in toxins toxin; cell-associated RTX toxin; vaccine production;
XX

```



PA (GENO-)INST GENOMIC RES.  
 XX  
 PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M,  
 PI Petersen J, Pizzo M, Rappunli R, Ratti G, Scalato E, Scarselli M,  
 PI Tettelin H, Venter JC;  
 XX  
 DR WPI: 2000-062150/05.  
 DR N-PSDB: AAZ54336.  
 XX  
 PT Novel Neisserial polypeptides predicted to be useful antigens for  
 PT vaccines and diagnostics  
 XX  
 PS Claim 2: Page 1244; 1453pp: English.  
 XX  
 CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941  
 CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides  
 CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent  
 CC PCR primers used in the exemplification of the present invention. The  
 CC polypeptides, the polynucleotides, antibodies and compositions of  
 CC the invention can be used as vaccines, as diagnostic reagents, and as  
 CC immunogenic compositions. The polypeptides can be used in the  
 CC manufacture of medicaments for treating or preventing infection due to  
 CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the  
 CC presence of Neisseria bacteria, or to raise antibodies. They may also  
 CC be used to screen for agonists or antagonists, which may themselves  
 CC have use as antibacterial agents. The polynucleotides of the invention  
 CC may also be used in gene therapy protocols.  
 XX  
 SQ Sequence 435 AA:  
 Query Match 12.1%, Score 261.5; DB 21; Length 435;  
 Best Local Similarity 25.2%; Pred. No. 3.9e-14;  
 Matches 102; Conservative 70; Mismatches 163; Indels 69; Gaps 14;  
 QY 23 IIALTFVGLGMAAARPLDSAVIANG-VYSAVVSQDQVHLECGMLAKILYRGEKVKAQO 81  
 Db 62 Imafa11a11-WSwtgkIdIvaasgktvsggrsktIqbletvvkvavhtdgqnvkqge 120  
 QY 82 VLEFEDPT-----QANAAGSITRNQOYA-LKAMEARLLAERDQPSISFPADLTG 130  
 Db 121 tlaeleavgtddvvgsegalgaqalskryeaavlaaestr-----tlv 163  
 QY 131 QRADPVAARAIADDEQAQFTERRTOTGGVDLMMAGRLQYQSEIEICIDKOTGKQO--- 186  
 Db 164 phldmagarslgsldadvgsaqvlabghyqawaagdaqgsalrghqaelgsakagckl 223  
 QY 187 --LGTIDE-LIDLKLVKDKGLVPPRLALAEARGS-----LSGSIGRTADRSAVQGA 239  
 Db 224 vsvga1eqgktadryrrlradnfinfisehafieqgsksvsnwndlestrgmrqigaalaae 283  
 QY 240 SPTOLKVRIOKEFFE---QVSQSTIEFVRRLAEYTEKEEVNASDAQRIKITYSPNGTAQ 296  
 Db 284 gnrvtlnqtlktdctldalrqaegldyrgq----tdk---akrgqgltmtdispsadgltvq 336  
 QY 297 NLRFTEGAVVRAAPLVADIAPDEAFVLIQAHFQPTDVONVHMGWTEVRLPAF----- 350  
 Db 337 elatytvgvgvaagqkmmvrapddkmdvexlvlkdkifveggqdaavkiesfpylryg 396  
 QY 351 HSAGNPDPERHDPVAVADRISPOKQARFLGCTIVRVKQOLPRH 394  
 Db 397 yltgkvksvshdavs-----heg-----lglvyltavvsldkh 428  
 RESULT 12  
 AAY24069  
 ID AAY24069 standard; peptide: 425 AA.  
 XX  
 AC AAY24069;  
 XX  
 DT 09-SEP-1999 (first entry)  
 XX  
 DE *Salmonella typhimurium* peptide sequence SEQ ID NO:38.



Db 299 qgltfiegrfyraltlrglvsqgfvekqadvidqrahal-----elqrerltladalaq 353  
OY 242 TOLKVRQIKKEFEEOVS-----QSITETRVRLAEVTEKEVVASDAQRKIKIVSPVNGTA 295  
Db 354 aqaelqqlvslrltqqlalagaaslgadrttaieqaaasrwevrapragr----- 401  
OY 296 QNRFFTEGAVVRAAEPLVDIAPEDEAFVIOAHFOPTVDNVHMGWTEVR---LPAFH- 351  
Db 402 valrlplrgqgavqgqladllpstatelvvlyapstraagllpglpglrfdaipyghy 461  
OY 352 -----SAGNDPERRHPPVAVADRISDPQOKARLFLGIVRDVQOLPRLRGRTAG 402  
Db 462 gqfagrvevlaaapepprad---aalaseplyrvvrla---gdaalragh-aavlrpg 513  
OY 403 MPAQVIVPTGERVLYOYLFSPFLRDTLRTMR 433  
Db 514 mrvqglalalewrtfswafepI-sllngtlir 543  
RESULT 14  
AAC98387  
ID AAC98387 standard; Protein: 355 AA.  
XX AAC98387:  
XX  
XX 21-SEP-2001 (first entry)  
XX  
XX Escherichia coli protein sequence SRQ ID NO:435.  
XX  
XX Escherichia coli; identification: proliferation; microorganism:  
XX antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis;  
XX bacterial growth inhibition.  
XX  
XX Escherichia coli.  
XX  
XX WC200148209-A2.  
XX  
XX 05-JUL-2001.  
XX  
XX 19-DEC-2000; 2000WO-US34419.  
XX  
XX 23-DEC-1999; 99US-0173005.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX PI Forsyth RA, Ohlsen KL, Zyskind JW;  
XX  
XX WPI: 2001.457376/49.  
XX  
XX N-PSDB: AAH81443.  
XX  
XX Novel nucleic acids encoding proteins required for Escherichia coli  
XX proliferation, useful for screening for antimicrobial agents -  
XX  
XX Claim 19, Page 554-555; 596pp; English.  
XX  
XX The present invention describes a purified or isolated nucleic acid  
XX sequence (I) consisting essentially of one of the 93 nucleotide sequences  
XX given in AAH81292 to AAH81294, where expression of the nucleic acid in a  
XX microorganism is capable of inhibiting proliferation of a microorganism.  
XX (I) have antibacterial and antibiotic activities, and can be used in  
XX gene therapy. Expression of (I) in a microorganism inhibits proliferation  
XX of the microorganism, and the manufactured antibiotic is useful for  
XX reducing the activity or level of a gene product required for  
XX proliferation of a microorganism in a subject, specifically humans. The  
XX nucleic acids that inhibit bacterial growth or proliferation can be used  
XX as antisense therapeutics for killing bacteria. In addition to  
XX therapeutic applications, the nucleic acid sequences complementary to  
XX sequences required for proliferation can be used as diagnostic tools.  
XX For example, nucleic acid probes complementary to proliferation-required  
XX sequences that are specific for particular species of microorganisms can  
XX be used as probes to identify particular microorganism species in  
XX clinical specimens. AAH81295 to AAH81487 encode the Escherichia coli

CC proteins given in AAC98239 to AAC98431, and AAH81488 to AAH81491  
CC represent oligonucleotides, which are used in the exemplification of the  
CC present invention.  
XX  
XX Sequence 355 AA;  
Query Match 6.6%; Score 142; DB 22; Length 355;  
Best Local Similarity 21.2%; Pred. No. 0.0006;  
Matches 87; Conservative 65; Mismatches 158; Indels 100; Gaps 20;  
OY 22 GIIATLFVGLGNAAFAPL---DSAVIANGVSAEVSQDYHLEGMLKILVREGEKVK 78  
Db 14 gllav--aaivawwllrpagypegfavngrlea-tevdaskiagrlidilvkegkfvr 70  
OY 79 AGCVLFEFLDPTQANNAAGITRNOYVALKAMEARLLAERDORPSISPADLTSORADPMVA 138  
Db 71 egevlakmd-----tr-----viquegrleaiqakeags----- 99  
OY 139 RAIDEOAOFTERKQRTIQGVVDLMNAORLQYOSFIEGI-DROTQGLK-DQLGFIEDELID 196  
Db 100 -avaaaqalllegqrsetraagslvn---grgaeldsvakhrtrsrjaqrgalsaqld 154  
OY 197 LRLKYDKGLVPRRLALLEARAGSLSGSIGRLTADRSAKAVOGASDTOLKVRQIKKEFFEQ 256  
Db 155 -----ddraaesarralesakagysaskaieartnliq----- 190  
OY 257 VSQSITETRVRLAEVTEKEVVASDAQRKIKIVSPVNGTAQNLRFTEGAVVRAAEPLVDI 316  
Db 191 -----aglrvaagaqaterliaaddidseik-aprdgrvq-yrvaepegvlaagrvllm 242  
OY 317 APEDEAFVIOAHFOPTD-VDNVHMG-----WTEVRLPA-----FHSAGNDPERRHD 363  
Db 243 v--disdymtflippteagqlkigearllldaapdlrtpatistfvaasqftrp---t 297  
OY 364 VAVADRISDPQOKARLFLGIVRDVQOLPRLRGRTAGMPAQVIVTGE 413  
Db 298 veld-----erlkimfvkarippellqghle-yvktglpgvawvrve 341  
RESULT 15  
AAU36320  
ID AAU36320 standard; Protein: 285 AA.  
XX  
XX AAU36320:  
XX  
XX 14-FEB-2002 (first entry)  
XX  
XX DT Pseudomonas aeruginosa cellular proliferation protein #310.  
XX  
XX DE  
XX  
XX Antisense; prokaryotic cellular proliferation protein;  
XX antibiotic; antibacterial; drug design.  
XX  
XX OS Pseudomonas aeruginosa.  
XX  
XX WC200170955-A2.  
XX  
XX 27-SEP-2001.  
XX  
XX 21-MAR-2001; 2001WO-US09180.  
XX  
XX 21-MAR-2000; 2000US-191078P.  
XX  
XX 23-MAY-2000; 2000US-206848P.  
XX  
XX 26-MAY-2000; 2000US-207727P.  
XX  
XX 23-OCT-2000; 2000US-242578P.  
XX  
XX 27-NOV-2000; 2000US-253625P.  
XX  
XX 22-DEC-2000; 2000US-257931P.  
XX  
XX 16-FEB-2001; 2001US-269308P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
XX Yamamoto RT, Xu HH;



XX WPI; 2001-611495/70.  
DR N-PSDB; AAS54179.  
XX  
XX  
XX New polynucleotides for the identification and development of  
XX antibiotics, comprise sequences of antisense nucleic acids -  
XX  
XX Example 3; Seq ID No 11913; 511pp; English.  
XX  
XX The invention relates to antisense inhibitors of genes essential to  
XX prokaryotic cellular proliferation, their use in identifying the  
XX genes, their use in the discovery of novel antibiotics, the essential  
XX genes themselves and the encoded proteins. The prokaryotes used are  
XX Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
XX pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
XX invention is also useful for the identification of potential new targets  
XX for antibiotic development. The antisense nucleic acids can also be used  
XX to identify proteins used in proliferation, to express these proteins,  
XX and to obtain antibodies capable of binding to the expressed proteins.  
XX The proteins can be used to screen compounds in rational drug discovery  
XX programmes. The antisense nucleic acid sequence is also useful to screen  
XX for homologous nucleic acids which are required for cell proliferation in  
XX a wide variety of organisms. The present sequence represents an  
XX essential prokaryotic cellular proliferation protein.  
XX Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic  
XX format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 285 AA:  
SQ

Query Match 6.0%; Score 130; DB 22; Length 285;  
Best Local Similarity 21.2%; Pred. No. 0.0048;  
Matches 87; Conservative 56; Mismatches 103; Indels 164; Gaps 20;  
QY 23 IIALTEVGLGMAFA-----PLDSAVIANGVSAEVSQDVQHLEGMIAKILVREGKVK 78  
DB 14 vaaavaglavwryymspwtrdarvadvv---vapdv---sgwvtdlevkdqyvk 66  
QY 79 AGOVLEFLDPTQANMAAGITRNQYVALKAMEARLLAERDQPSISFPADLTSGRADPMVA 138  
DB 67 vgdvllmrid-----qeryqanleq-----a 86  
QY 139 RA1ADBEAQOTERRQTIQGOVDLMNAQRLQYOSIEIDRQTOGLKDQGFIEDELIDLR 198  
DB 87 ravaethq-----qy-----llrqneaar----- 106  
QY 199 KLYDKGLVPRRLALEARAGSLSGISRLTADRSKAVOGASPTOLKVRQIKOEFPOVS 258  
DB 107 -----tsrl-----qfgalsaeke-----naqlnaalareygeaia 139  
QY 259 QSTETFRVRLAEVTEKEVVASDAOKRIKIVSPVNGTAQNLR-----FTEGAVVRAAEPL 313  
DB 140 q-----vklaeln-----lkreslraarngvtnrlagmyatagqawma---l 181  
QY 314 VDIAPEDAEAVIQAHQPTVDVNHMGMTVEVRLPAFHSAGNPDERHDVA--VADRIS 371  
DB 182 vd-----qgstfyvayfeetklpglrvgmraqvrl---msgdqpIdgtlvesissgltDns 234  
QY 372 DPQKQ-----ARLFLGI-VRVDVVKQLPRHLRGRTAGMPAQIV 409  
DB 235 tpdgqlanveptfnwrlqgrlprvrlrlqypradv--hlsagmtasvly 282

Search completed: July 19, 2002, 10:43:19  
Job time: 139 sec

**THIS PAGE BLANK (USPTO)**



Db 5 IGEPODSTYETEIPQDERFRFTMGVLVYIGLFGFLAANAAPLADKGVASGCVTVSGNRK 64  
QY 57 DVOHLEGMLAKLIVREGGEYKAGOVLEFEDPTQANAAGITRNOYVALKAMEARLLAER 116  
Db 65 TVAPASGIITKINNAVRGDKVAKAGEVVLVSQVQAQOVSLNDQYTTTLATBGRLLAER 124  
QY 117 DORPSISFPADLTQSARADPMVARAIADEQAQFTERRQTIQOVUDLMAARLOYSIEGI 176  
Db 125 DGLSIVFSPFLDAVKKPRVAETIALQTLQFASRQALQSEID-----GKQSDGI 177  
QY 177 DROTQGLKD-----OLGFIEDLIDRLKYDKGLVPRRLALEARASLSGSGIRLT 229  
Db 178 RFOLKGLQDSRGKNOIQLSSLRQOMNSMKOLADGYLPRNRKYLEVORQFAEVNSSIDETV 237  
QY 230 ADRSKAVQASDPTQLKVRQIKQEFFEQVQSITETRYRLAEVTEKEVASDAQKRIKIVS 289  
Db 238 GRIGQLQKLLSQQRIDQRFADYQREVRLQATQMDASEFRKLLQMDPDLGNTAITS 297  
QY 290 PVNGTAONLFEETEGAVVRAEPLVDIAPEDEAFVIOAHFOPTDVNVHMGMTVEVRLPA 349  
Db 298 PVDGTAVGLNFTGGVVGAGDHLMDVVPQSQATLVDSRLKVDLDFKYYNGLPVDLMFTA 357  
QY 350 FHSANGNDPERHDPVAVADRISD-----PQKQARLFL---GIVAVDYKOLPRLHGRVTA 401  
Db 358 FNOKTKPRIGTVTLVSADRLVDKANGEPYQOMQVTVSPGKMLSGEDIKP----- 409  
QY 402 GMPAOVIVPTGERTVLOYLFSPRLDRLTTMREE 435  
Db 410 GMPVEVFKTGSRLSLTYLFRPILIDRAHTSLTEE 443

## RESULT 2

US-08-620-605D-3  
Sequence 3, Application US/08620605D  
Patent No. 5846811

## GENERAL INFORMATION:

APPLICANT: SHIBATANI, TAKEJI  
APPLICANT: AKATSUKA, HIROYUKI

APPLICANT: KAWAI, ERI  
TITLE OF INVENTION: GENE PARTICIPATING IN THE MECHANISM OF

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH

STREET: PO BOX 747  
CITY: FALLS CHURCH

STATE: VA  
COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/620,605D  
FILING DATE: 22-MAR-1996

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: SVENSSON, LEONARD R  
REGISTRATION NUMBER: 30,330

REFERENCE/DOCKET NUMBER: 0020-3955  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)-205-8000  
TELEFAX: (703)-205-8050

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

LENGTH: 443 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-08-620-605D-3

Query Match 23.6%; Score 510.5; DB 2; Length 443;  
Best Local Similarity 29.1%; Pred. No. 1.7e-40;  
Matches 132; Conservative 88; Mismatches 195; Indels 39; Gaps 7;  
QY 6 IORPTNE-----QAVARIGYGIATLFFVGLLGMAAPPLDSAVIANGVSAEVS-Q 56  
Db 5 IGEPODSTYETEIPQDERFRFTMGVLVYIGLFGFLAANAAPLADKGVASGCVTVSGNRK 64  
QY 57 DVOHLEGMLAKLIVREGGEYKAGOVLEFEDPTQANAAGITRNOYVALKAMEARLLAER 116  
Db 65 TVAPASGIITKINNAVRGDKVAKAGEVVLVSQVQAQOVSLNDQYTTTLATBGRLLAER 124  
QY 117 DORPSISFPADLTQSARADPMVARAIADEQAQFTERRQTIQOVUDLMAARLOYSIEGI 176  
Db 125 DGLSIVFSPFLDAVKKPRVAETIALQTLQFASRQALQSEID-----GKQSDGI 177  
QY 177 DROTQGLKD-----OLGFIEDLIDRLKYDKGLVPRRLALEARASLSGSGIRLT 229  
Db 178 RFOLKGLQDSRGKNOIQLSSLRQOMNSMKOLADGYLPRNRKYLEVORQFAEVNSSIDETV 237  
QY 230 ADRSKAVQASDPTQLKVRQIKQEFFEQVQSITETRYRLAEVTEKEVASDAQKRIKIVS 289  
Db 238 GRIGQLQKLLSQQRIDQRFADYQREVRLQATQMDASEFRKLLQMDPDLGNTAITS 297  
QY 290 PVNGTAONLFEETEGAVVRAEPLVDIAPEDEAFVIOAHFOPTDVNVHMGMTVEVRLPA 349  
Db 298 PVDGTAVGLNFTGGVVGAGDHLMDVVPQSQATLVDSRLKVDLDFKYYNGLPVDLMFTA 357  
QY 350 FHSANGNDPERHDPVAVADRISD-----PQKQARLFL---GIVAVDYKOLPRLHGRVTA 401  
Db 358 FNOKTKPRIGTVTLVSADRLVDKANGEPYQOMQVTVSPGKMLSGEDIKP----- 409  
QY 402 GMPAOVIVPTGERTVLOYLFSPRLDRLTTMREE 435  
Db 410 GMPVEVFKTGSRLSLTYLFRPILIDRAHTSLTEE 443

## RESULT 3

US-08-772-270A-13  
Sequence 13, Application US/08772270A  
Patent No. 6019984

## GENERAL INFORMATION:

APPLICANT: MacInnes, Janet  
APPLICANT: Ricciatti, Paul

APPLICANT: Mallard, Bonnie  
APPLICANT: Rosendal, Soren

TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR  
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bereskin & Parr

STREET: 40 King Street West  
CITY: Toronto

STATE: Ontario  
COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/772,270A  
FILING DATE: December 23, 1996

CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:

NAME: Gravelle, Michelle  
REGISTRATION NUMBER: 40,261

REFERENCE/DOCKET NUMBER: 6580-81  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311

```

; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
; US-08-772-270A-13

```

```

Query Match      14.5%; Score 313.5; DB 3; Length 477;
Best Local Similarity 25.3%; Pred. No. 1.1e-21;
Matches 112; Conservative 88; Mismatches 190; Indels 53; Gaps 13;

```

```

QY 19 ICGIATLTFVGLGMAAPPLDSAVIANG-VSAEVSQDVQHLEGMLAKILVREGKV 77
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 60 IAYLMFLFLALV-ISIISKVEIVASATGKLVFSGHSEIKPIENALVKDIEVKDQFV 118
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 78 KAGOVLEL-----DPTQANAAAGITRNQYVALKAMEARLLAERDPSISF-PADLT 129
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 119 EKQDLHLTLALGADADQOKTKVSLGERLDGYRKSLYSI--EHNRLPLDPMQADFD 176
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 130 S-ORADPMVARAIAD-----QAQFERRQTIQGVDMNAQRLQYSEIEGIDR 178
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 177 SVGEOKTGARHLITEGFETWQOKYOKELAYQKQKOTVLANIRKYSASNIE----- 232
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 179 QTGGLKDLQGLFIEDELIDRLKLYDKGLVPRRLALAEAR---AGSISGIGRLTADRSK 234
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 233 -----KEKLSLDLKKLYDVSISKHELAEQENRYEASNELSYQSHLKEVESD 280
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 235 AVOGASDTQLKVRQIKQEFEEYOSITETRVRLAEVTEKEVNASDQKRKIY--IYSPVN 292
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 281 LIAQEDLKVLTOLFKSDILEKLOQNOIREKQLTLELKN-----QROLASITIRPVS 334
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 293 GTAONLREFTEGAVVRAEPLVDIAPEDEAFVIOAHFQPTVDNVHMGVTEVRLPAFHS 352
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 335 GTVQOLKTHGKGVVTTAEFLMVAIPEDDVLEVSALQONKDVGFELGQDAVIVKEFFPY 394
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 353 AGNPPERHDPVAVADRIPOKQARLFLGIVRVYDKQLPRLHGR--RVTAGMPAQVIVPT 411
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 395 TRYGYLGVKKTITLLDAIEHPQ--LGIVFNSIIEINKTITDGDKEIQLGSGMSVIAEIKT 453
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 412 GERTVLOYLESPLRDLRTTMR 434
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 454 GERSVLSFLSPLSESTESLRE 476
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 4
; US-08-772-270A-5
; Sequence 5, Application US/08772270A
; Patent No. 6019984
; GENERAL INFORMATION:
; APPLICANT: Macinnes, Janet
; APPLICANT: Ricciattl, Paul
; APPLICANT: Mallard, Bonnie
; APPLICANT: Rosendal, Soren
; TITLE OF INVENTION: NOVEL BACTERIAL PREPARATIONS, METHOD FOR
; TITLE OF INVENTION: PRODUCING SAME, AND THEIR USE AS VACCINES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/772,270A
; FILING DATE: December 23, 1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 6580-81
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 478 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Actinobacillus pleuropneumoniae
; US-08-772-270A-5

```

```

Query Match      12.1%; Score 262; DB 3; Length 478;
Best Local Similarity 24.2%; Pred. No. 9.4e-17;
Matches 111; Conservative 86; Mismatches 178; Indels 84; Gaps 18;

```

```

QY 19 ICGIATLTFVGLGMAAPPLDSAVIANGVSAEVSQDVQHLEGMLAKILVREGKV 77
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 60 IAYLMFLFLALV-ISIIVHEIVATATGKLAFSDRSKEIKPIENALVKEIVQDQFV 118
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 78 KAGOVLELPTQANAAAGITRNQYVALKAMEAR--LLAE--RDQPSISFAD--LTS 130
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 119 EKQDLHLTLALGADADQOKTKSSLTLEKRYRELLERAVAADRLPLETKDEFKHA 178
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 131 QRADPMVARAIADQOAFTR-----ROTIQGVDMNAQRLQYSEIEGIDRQT 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 179 TEEDKTRIRRLITEGFQFAMQKQYOKELALQREAEKQVLANIRK-----YEGISR-- 230
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 QGLKDLQGLFIEDE-LIDRLKLYDKGLVPRRLALAEAR---AGSISGIGRLTADRSKA 235
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 231 -----VENERKLQDKLKLKFNKSKTSKHDVLTQENRHEAVNEIAYKSRINEVESDL 281
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 236 VOGASDTQLKVRQIKQEFEEYOSITETRVRLAEVTEKEVNASDQKRKIY--SPVNG 293
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 282 RQKKEEHLITOLFRADILEKLONV-EAKQSLLELEK-----NEQRLASVIRAFVSG 335
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 294 TAONLREFTEGAVVRAEPLVDIAPEDEAFVIOAHFQPTVDNVHMGVTEVRLPAFHSA 353
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 336 TVQOLKTHGKGVVTTAEFLMVAIPEDDVLEVTALQNKDIGFLVGQDAVIVKEFFPY 392
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 354 GNPDPERHDPVA-----VADRIPOKQARLFLGIVRVYDKQLPRLHGR--- 398
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 393 -----PYTRYGYLMGVKNITTEAIEHPQ--LGIVFNSIISIDRKT-----LSGDKG 438
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 399 ---VTAGMPAQVIVPGEPTVLOYLESPLRDLRTTMR 434
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 439 EIELGMSVTAIEIKGERSVISTLSPLESVSESLRE 477
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 5
; US-08-853-659A-38
; Sequence 38, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saifer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods of Detection Of Salmonella
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman

```

```

ADDRESS: Intellectual Property Services
ADDRESS: Battelle Memorial Institute
ADDRESS: PNNL P.O. Box 999
STREET: Washington Way
CITY: Richland
STATE: Washington
COUNTRY: U.S.A.
ZIP: 99352
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
COMPUTER: IBM PC/XT/AT
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Processor (Wordperfect 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,659A
FILING DATE: Unknown
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: none
FILING DATE: n/a
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 425 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-853-659A-38

```

```

Query Match          9.3%; Score 201.5; DB 2; Length 425;
Best Local Similarity 21.0%; Pred. No. 4.6e-11;
Matches 94; Conservative 93; Mismatches 187; Indels 73; Gaps 19;

```

```

QY 7 ORPDNFOAVARIGYITALTPEVGLGMAAFAPLDSAVIANGVSA-EVSODVOHLCGM 65
DB 3 RROSDHLMIT-----IISLTIL-IILTYFIEINSYHVGQVITTDNMQILISKSGT 55
QY 66 LAKLIVREGKRVAGQVL---FELDPTQANAAAGITRNOYVALKAMEARLLAERDQRPST 122
DB 56 IODIYVAEGDTVKKGELLAKVND-----LQKEYQRYRTQKGYL--DKDVN-EI 102
QY 123 SFPRDLTSQRADPNV--ARAIADQAOFTERRQTIQGOVDMNQRIQOYSEITGIDRQT 180
DB 103 SFILDKNEGSLTLDGTRISLNREK--ANIELVHSQIRAKELKYSLSSEISGLQEKL 160
QY 181 OGKDOGFIEDELIDIRKLYDKGLV-----RPRLALEARAGSLSGISGRLTADRS 233
DB 161 SKRKEALALAEENILSPYKKGISPYTNFLNKKQAVIKYKSEINDIESTI-TLKKDDI 219
QY 234 KAVOGASDTOLKVRQIKOEFEFOVSQSIETETRV--RLAEVTEKEVVASDAQRIKIVSPV 291
DB 220 ELV--VNDIEALNNEIRLSKITSKLNQLELVVNSTLKVLEKQINNEED-----IYSPV 271
QY 292 NGTQO--NLRFTEGAVVRAEPLVDIAPDEAVIQAHQPTVDVNVHGMVTEVRLPA 349
DB 272 DGVLYRKINKSATTHGVYIQADLLFEIKPKVRYMLADVKKILPKYRDQIYDEAVKLDVQS 331
QY 350 FHSAGNPDPERHDVAADRISDPQKQ-----ARLEFIVRVDVQO-----LPHLRG 397
DB 332 I-----IQPKTKSNATINDISPSYEENGTIGQRYKVIILADVNNEDDLRLKLP----- 382
QY 398 RVTAGMPAQVIVPTGERTVQLYLFSP 424
DB 383 ---GMTVDASVITGKSHIMEYLLSPL 405

```

```

RESULT 6
US-08-868-699A-2
; Sequence 2, Application US/08868699A
; Patent No. 6204019
; GENERAL INFORMATION:
; APPLICANT: O'Dwyer, Karen
; APPLICANT: Perry, Carolanne
; APPLICANT: Warren, Richard L.

```

```

TITLE OF INVENTION: No. 6204019e1 Compounds
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/868,699A
FILING DATE: 04-JUN-1997
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10012
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 796 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-868-699A-2

```

```

Query Match          5.5%; Score 118.5; DB 4; Length 796;
Best Local Similarity 20.6%; Pred. No. 0.0099;
Matches 63; Conservative 52; Mismatches 120; Indels 71; Gaps 11;

```

```

QY 10 TDFNFOAVARIGYITALTPEVGLGMAAFAPLDSAVIANGVSA-EVSODVOHL----- 61
DB 358 TTFQNLFKFESPSGMTATGKLGSESEFPLYSKTI---VVOAPTDKAIQRIDEPDKVFR 413
QY 62 -----EGMKLAKILVREGKRV--AGQVLFELDPTQANAAAGITRNOYV 103
DB 414 SVDEKNTAMTHDIVELHETGEPVLLITRTADAEVSEVLFQMD-----IPNNLLI 464
QY 104 ALK-AMBARLLAERDQRPSTSFADLTSSORADPNVVARAIADQAOFTERRQTIQGOVDLM 162
DB 465 AQNVAKKEQMAIEAGQIGSMVTATSMAGRGTDILGEGV-----EALGGIIVII 513
QY 163 NAQRLQYQSEIEGIDROTQGIKQDLG-----FIEDELIDIRKLYDKGLVPRPPLA 213
DB 514 HE-----HMENSRVDRDLRGRSGRGDPGSSCIYISLDIYLV--KRWSDSNLENNQLYS 566
QY 214 LEARAGSLSGISGRLTDRSKAVOGASDTQ--LKVRQIKOEFEFOVS--QSTIETVRILA 269
DB 567 LDAQRLSQSNLENKRVQKIYVKAQRISEEGVAKREMANEFERSISIQRDLYTEERRVL 626
QY 270 EVTEKE 275
DB 627 EIDDAE 632

```

```

RESULT 7
US-09-757-014-2
; Sequence 2, Application US/09757014
; Patent No. 6348342
; GENERAL INFORMATION:
; APPLICANT: O'Dwyer, Karen

```

1 Perry, Caroline  
 2 Warren, Richard L.  
 3  
 4 TITLE OF INVENTION: NO. 6348342el Compounds  
 5  
 6 NUMBER OF SEQUENCES: 6  
 7  
 8 CORRESPONDENCE ADDRESS:  
 9 ADDRESSEE: Dechert, Price & Rhoads  
 10 STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
 11 City: Philadelphia  
 12  
 13 STATE: PA  
 14  
 15 COUNTRY: USA  
 16  
 17 ZIP: 19103-2793  
 18  
 19 COMPUTER READABLE FORM:  
 20  
 21 MEDIUM TYPE: Diskette  
 22  
 23 COMPUTER: IBM Compatible  
 24 OPERATING SYSTEM: Windows  
 25 SOFTWARE: FASTSEQ for Windows Version 2.0b  
 26  
 27 CURRENT APPLICATION DATA:  
 28 APPLICATION NUMBER: US/09/757,014  
 29 FILING DATE: 09-Jan-2001  
 30  
 31 CLASSIFICATION: <Unknown>  
 32  
 33 PRIOR APPLICATION DATA:  
 34 APPLICATION NUMBER: 08/868,699  
 35 FILING DATE: <Unknown>  
 36  
 37 ATTORNEY/AGENT INFORMATION:  
 38 NAME: Falk, Stephen T  
 39  
 40 REGISTRATION NUMBER: 36,795  
 41  
 42 REFERENCE/DOCKET NUMBER: GML0012  
 43  
 44 TELECOMMUNICATION INFORMATION:  
 45 TELEPHONE: 215-994-2488  
 46  
 47 TELEFAX: 215-994-2222  
 48  
 49 TELEX: <Unknown>  
 50  
 51 INFORMATION FOR SEQ ID NO: 2:  
 52  
 53 SEQUENCE CHARACTERISTICS:  
 54 LENGTH: 796 amino acids  
 55 TYPE: amino acid  
 56 STRANDEDNESS: single  
 57  
 58 TOPOLOGY: linear  
 59  
 60 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 61  
 62 US-09-757-014-2

```

Query Match          5.5%: Score 118.5: DB 4: Length 796:
Best Local Similarity 20.6%: Pred. No. 0.0099:
Matches 63: Conservative 52: Mismatches 120: Indels 71: Gaps 11.

QY      10 TDNFQAVARIGYGIATLFTVGLLGMAAEPAPLDSAVIANGVYSAEVSODVQH----- 61
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      358 TTFEONLFKEEFSEFGMATGKLGSEFFDYLSKI-----VVOAPMDKAIORIDEPPKVR 413
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      62 -----BGMGLAKLIVVEGEKVK--AGOVLPFLDPTQNAAGITRNQVY 103
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      414 SVDEKNINMIDIVLHETGRVLLITTAEEAEKFESEVLFQMD-----IPNNLLI 464
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      104 ALK-AMEARLLAERDQRPISFPADLTISQARDPMYARAIADBOAFTERQRTIOGOVDM 162
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      465 AONVAKEAOMIAEAGQISMTVATSMAGRTDIKLGEGV-----EALGGLAVII 513
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      163 NAQRLOVQSEIEGIDPRTQGLKLDQ-----FIEDELIDLRKLYDKGVPRPRLA 213
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      514 HE-----HMENSRVDRQLRGSRGRGDDGSCSITYSLDYLV--KRMSDNLAEENQLYS 566
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      214 LEARAGLSGSIIGRLTADRSKAVOGASPTQ-LKVRQIKOEFEQVYS--QSITETRVRLA 269
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      567 LDAOGLTSSQSNLFNRVKQIVVKAQRISSEOGVKAREMANEEFKSISIORDLVEERNRVL 626
      | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      270 EVTEKE 275
      | : : |
Db      627 EIDDAE 632

```

Patent NO. 56235879  
GENERAL INFORMATION:  
APPLICANT: Hayden, Michael R.  
APPLICANT: Hackam, Abigail  
APPLICANT: Huq, A.H.M. Mahbubul  
APPLICANT: Chopra, Vikramjit Singh  
APPLICANT: Kalschman, Michael  
TITLE OF INVENTION: Apoptosis Modulators That Interact with the  
TITLE OF INVENTION: Huntington's Disease Gene  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Oppedahl & Larson  
STREET: PO Box 5270  
CITY: Frisco  
STATE: CO  
COUNTRY: USA  
ZIP: 80443-5270  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS DOS 5.0  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/085,199B  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Larson, Marina T.  
REGISTRATION NUMBER: 32038  
REFERENCE/DOCKET NUMBER: UBC-P-013052  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (970) 668-2050  
TELEFAX: (970) 668-2052  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1090  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: no  
ORIGINAL SOURCE:  
ORGANISM: human  
FEATURE:  
OTHER INFORMATION: Huntington-interacting protein  
US-09-085-199B-5

```

Query Match          5.3%: Score 115: DB 4: Length 1090:
Best Local Similarity 20.1%: Pred. No. 0.034:
Matches 82: Conservative 72: Mismatches 131: Indels 122: Gaps 20:

QY 47 NGVYSAEVSODVDHLE--GGMLANILVREGKRYAGVLFELDPDTQANAAGITRNQYV 103
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 NGVNKDEKDHILERYLREISGLKQL--ENMKTESGRVVLQDKGVSELEADLAEQDHL 475

QY 104 ALKAME--ARLLAERDQRRSIFRADLTSGRADPVAALADE---QAQFTFR----- 152
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 476 RQQAADDCGFELRAELD-----ELRRQRDETEKQRSLEIRKKAQANQRYSLKE 526

QY 153 ---STIQGVOLM--NAQ-----RLQVQSEISGLDRTQGLKDLQGLTFED-----ELI 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 527 KYSELVQNHADLLRKNAEVLTKQVSNARAQVD--LEREKKELEDLSLERISDQGRKTOEOL 585

QY 196 DLRLKYDKGLVPRRLALEARAGSLGS-----IGRLTRDSKAVQAGASDTOLK 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 586 EYLESLKQELATSOR--DLOVQGLSLETSQAQSEANMAAEFALEKERPSILVGAHREEE 643

QY 246 VROIKQOEFEQVQSITETVRLAEVTEKEVVASQDAOKRIKIVSVNQTAOVLRFETGSA 305
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 644 LSAIRKE-----LDPTQLKLASTEESKCOLADQDKRMILV-----GSRKAEDVTDPA 691

QY 306 VVRAAE--PLVDIA-----PED-----EAFVQAQHF 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 692 INOLEEPPILISAGSADHLLSTVTSISSCIEQLEKSMYSOYLACPEDISGLHSTITLHL 751  
QY 330 QPTDVNVHMGWTEVRLPAFHSAGNDPDRHPDVAADRISDPQK 376  
Db 752 TS---DAIAGATTCLRAP-----PEP-----ADSLTEACKQ 780

RESULT 9  
US-09-085-199B-4  
; Sequence 4, Application US/09085199B  
; Patent No. 6235879  
; GENERAL INFORMATION:  
; APPLICANT: Hayden, Michael R.  
; APPLICANT: Hackam, Adigail  
; APPLICANT: Hug, A.H.M. Mahbulul  
; APPLICANT: Chopra, Vikramjit Singh  
; APPLICANT: Kalchman, Michael  
; TITLE OF INVENTION: Apoptosis Modulators That Interact with the  
; TITLE OF INVENTION: Huntington's Disease Gene  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Opedahl & Larson  
; STREET: PO Box 5270  
; CITY: Ftisco  
; STATE: CO  
; COUNTRY: USA  
; ZIP: 80443-5270  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS DOS 5.0  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/085,199B  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Larson, Marina T.  
; REGISTRATION NUMBER: 32038  
; REFERENCE/DOCKET NUMBER: UBC.P-013052  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (970) 668-2050  
; TELEFAX: (970) 668-2052  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 914  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: no  
; ORIGINAL SOURCE:  
; ORGANISM: human  
; FEATURE: Huntington-Interacting protein  
US-09-085-199B-4

Query Match 5.2%; Score 112; DB 4; Length 914;  
Best Local Similarity 20.1%; Pred. No. 0.051;  
Matches 82; Conservative 72; Mismatches 131; Indels 122; Gaps 20;

QY 47 NGVVAEVSQDVQHLE---GMLAKILVREGGEKYKAGCVLEFELDPQTQANAAGITRNOYV 103  
Db 243 NGVVKDEKHILERLYREISGLKAOI---ENMKTESQHVVLQIKGHVSELADLAEOQHL 299  
QY 104 ALKAME--ARLLARDORPSISFPADLTSGRADPMVAAIAD---QOQTFERR----- 152  
Db 300 ROQAAADCEFLRAELD-----ELRRQREDETERKARSISEIERKQANEQORSKLUKE 350  
QY 153 ---TIGOGVILM--NAO-----RLQYSEIEGIDRQTOGLKDOLGFTIED-----ELI 195  
Db 351 KYSELYVQNHADILKRNNAEVNTQVSMARQAYD--LEREKKELEDSEIERKSDGQORTQSOQL 409  
QY 196 DLRLYDKGLVPRRLALLEARAGSLSGS-----IGRLADRSKAVOGASDTOLK 245

Db 410 EYLESLKQELGTSQR--ELQVLOGSLETSAQSEANMAAEFAELEKERDLSVSGAHRNEE 467  
QY 246 VROIKQEFFEPOVSOSINETEVRRLAEVTEKEFVVASDAQKRIKIVSPVNGTAONIRFFTEGA 305  
Db 468 LSAIRKE-----LQDTQIKLASTEESMCQIAKQDKRMILV---GSKRAAEQVYLQDA 515  
QY 306 VVRAAE-PLVDAI-----PED-----EAFVIOAHF 329  
Db 516 INOLEEPPILISAGSADHLLSTVTSISSCIEQLEKSMYSOYLACPEDISGLHSTITLHL 575  
QY 330 QPTDVNVHMGWTEVRLPAFHSAGNDPDRHPDVAADRISDPQK 376  
Db 576 TS---DAIAGATTCLRAP-----PEP-----ADSLTEACKQ 604

RESULT 10  
US-08-642-846-2  
; Sequence 2, Application US/08642846  
; Patent No. 5886151  
; GENERAL INFORMATION:  
; APPLICANT: HOSTETTER, MARGARET K.  
; APPLICANT: GALE, CHERYL A.  
; APPLICANT: BENDEL, CATHERINE M.  
; APPLICANT: TAO, NIAN-JUN  
; APPLICANT: KENDRICK, KATHLEEN  
; TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE  
; TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MUEITING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.  
; STREET: 119 NORTH FOURTH STREET, SUITE 203  
; CITY: MINNEAPOLIS  
; STATE: MINNESOTA  
; COUNTRY: USA  
; ZIP: 55401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/642,846  
; FILING DATE: 03-MAY-1996  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MUEITING, ANN M.  
; REGISTRATION NUMBER: 33,977  
; REFERENCE/DOCKET NUMBER: 110.00280101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-305-1217  
; TELEFAX: 612-305-1228  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1664 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-642-846-2

Query Match 5.0%; Score 109; DB 2; Length 1664;  
Best Local Similarity 19.5%; Pred. No. 0.24;  
Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;

QY 92 NAAAGITRNOYVALKAMEARLLARDORPSISFPADLTSGRADPMVAAIADDEQAQTFER 151  
Db 974 NAKKGVTDQDEYI-----NAKLVQDKPKKNSI-----VTDPRDREYEL 1010  
QY 152 RQT-----IQGVLDLMAAQRLQYSEIEGIDRQTOGLKDOLGFTIEDLIDLRK- 199  
Db 1011 QQTASINHATIDSSIVGRPDSISTDMRLPYLSD--ELKKRPRTALLSADRILFMQEVNHPILRSN 1069



QY 200 -----LYDKGLVPRRLALLERAGSLSGSIGRLTADRSKAVOGASDTOLKYROI 249  
Db 1070 SVLVHPGAGATNSMLPEPDFELINSPARVNSNDNV-----AISGNAST-ISTNQL 1122  
QY 250 KQEFEEQ-VQSQSTE-----TRVRLAEVTEKEVVASDAQ 282  
Db 1123 DMNFDQATIGQIOEOPASKSANTVTRGDDGLASABETPTPTPKKSSISSKPAKLSSAS 1182  
QY 283 KR--IKIVSPVNGTAONLRFETEGAVVRAEPLVDIAREDEAVIQAHQPTVDN--V 337  
Db 1183 PRKSPIKIGSPV-----RVIKKNGSIAGIEPIPKATHKP-----KKSFOGNEISNHKV 1230  
QY 338 HMGVTE-----VRLPAFHSAGN--PDP-----ERHD-- 362  
Db 1231 RDGGISPSGSEHQHNPMSVSPSOYTDATSTVPDENKDVQHKPREKOKKHHNRHHN 1290  
QY 363 -----PAVADRISDP--QKQARLF--LGIVRVYKOLPRHLGRYT 400  
Db 1291 HHKQKTDIPGVVDEIPDVGLQERGLKLFVRLGIKINILPDIINH-KGRFT 1340

RESULT 11  
US-09-264-604-2  
; Sequence 2, Application US/09264604  
; Patent No. 6346411  
; GENERAL INFORMATION:  
; APPLICANT: HOSTETTER, MARGARET K.  
; APPLICANT: GALE, CHERYL A.  
; APPLICANT: BENDEL, CATHERINE M.  
; APPLICANT: TAO, MIAN-JUN  
; APPLICANT: KENDRICK, KATHLEEN  
; TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE  
; TITLE OF INVENTION: PROTEIN, ANTIBODIES, AND METHODS OF USE  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MUETING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.  
; STREET: 119 NORTH FOURTH STREET, SUITE 203  
; CITY: MINNEAPOLIS  
; STATE: MINNESOTA  
; COUNTRY: USA  
; ZIP: 55401  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/264,604  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/642,846  
; FILING DATE: 03-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MUETING, ANN M.  
; REGISTRATION NUMBER: 33,977  
; REFERENCE/DOCKET NUMBER: 110,00280101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-305-1217  
; TELEFAX: 612-305-1228  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1664 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-264-604-2

Query Match 5.08; Score 109; DB 4; Length 1664;  
Best Local Similarity 19.58; Pred. No. 0.24;

Matches 80; Conservative 61; Mismatches 124; Indels 146; Gaps 20;  
QY 92 NAAGTIRNOGVALKANEARLLARDQRPSTSPFADLTSGRADPMVARALINDEQAQFTER 151  
Db 974 NAKGVATQDEYT-----NAKLVDQKPKKNST-----VTPDPRVEEL 1010  
QY 152 ROT-----IGQVDMANARLOYOSEIGIDRQTOGLKDQGFIEDELIDRK- 199  
Db 1011 QQTASINNATIDSSITYGKPDISTDMYRLYSD-ELKKRPRTLLSADRLFMQGEVHPLRSN 1069  
QY 200 -----LYDKGLVPRRLALLERAGSLSGSIGRLTADRSKAVOGASDTOLKYROI 249  
Db 1070 SVLVHPGAGATNSMLPEPDFELINSPARVNSNDNV-----AISGNAST-ISTNQL 1122  
QY 250 KQEFEEQ-VQSQSTE-----TRVRLAEVTEKEVVASDAQ 282  
Db 1123 DMNFDQATIGQIOEOPASKSANTVTRGDDGLASABETPTPTPKKSSISSKPAKLSSAS 1182  
QY 283 KR--IKIVSPVNGTAONLRFETEGAVVRAEPLVDIAREDEAVIQAHQPTVDN--V 337  
Db 1183 PRKSPIKIGSPV-----RVIKKNGSIAGIEPIPKATHKP-----KKSFOGNEISNHKV 1230  
QY 338 HMGVTE-----VRLPAFHSAGN--PDP-----ERHD-- 362  
Db 1231 RDGGISPSGSEHQHNPMSVSPSOYTDATSTVPDENKDVQHKPREKOKKHHNRHHN 1290  
QY 363 -----PAVADRISDP--QKQARLF--LGIVRVYKOLPRHLGRYT 400  
Db 1291 HHKQKTDIPGVVDEIPDVGLQERGLKLFVRLGIKINILPDIINH-KGRFT 1340

RESULT 12  
US-08-816-693A-51  
; Sequence 51, Application US/08816693A  
; Patent No. 5874241  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S  
; APPLICANT: Turek, Fred W  
; APPLICANT: Pinto, Lawrence H  
; TITLE OF INVENTION: Clock Gene and Gene Product  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/816,693A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5874241thrup, Thomas E  
; REGISTRATION NUMBER: 33,268  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 747 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-816-693A-51



Search completed: July 19, 2002, 10:41:39  
 Job time: 39 sec

```

RESULT 15
US-08-179-738-5
: Sequence 5, Application US/08179738
: Patent No. 5578462
: GENERAL INFORMATION:
: APPLICANT: Seizinger, Bernd R.
: APPLICANT: Kley, Nikolai A.
: APPLICANT: Bianchi, Albert B.
: TITLE OF INVENTION: No. 5578462el NF2 Isoforms
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Reed & Robins
: STREET: 635 Bryant Street
: CITY: Palo Alto
: STATE: California
: COUNTRY: U.S.A
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/179,738
: FILING DATE: 10-JAN-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Robins, Roberta L.
: REGISTRATION NUMBER: 33,208
: REFERENCE/DOCKET NUMBER: 5998-0017
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 617-8999
: TELEFAX: (415) 327-3231
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 591 amino acids
: TYPE: amino acid
: TOPOLOGY: 1linear
: MOLECULE TYPE: protein
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: murine
: US-08-179-738-5
  
```

Query Match 4.9%; Score 106.5; DB 1; Length 591;  
 Best Local Similarity 21.6%; Pred. No. 0.087;  
 Matches 69; Conservative 54; Mismatches 117; Indels 79; Gaps 14;

```

QY 124 FPADLISQARADPVVARAIADDEQAQFTERRQTIQGYDLMNAQ---RLQYQSEIEGIDRQ 179
DB 283 FKFDSSKLRNKLILLOLCIGNHDLFMRRKRADSLEVOQMAQAREEKARQMERQRLARE 342
QY 180 TQGLKDLQGFIEDEL-IDLRKLYDKGLVPPRRLALEARAGSLSGISGRITADRSKAV-Q 237
DB 343 KQ-MREAEERTRELRLRLQMKFEATMANEALMRSEETADLAER-AQITTEBAKILAQ 400
QY 238 GASDTQKVKYQIQKEFFEOVSOSITETRVRLAE--VTEKEV---VASDAQKRIKIVSPV 291
DB 401 KAAEAQCEQMORIKATAIR-----TEEEKRLMQKYLEAEVLALNMAEESERRAKREADQ 454
QY 292 NGTAQNLRFTEGAVVRAAEPLVIDA-----PEDEAFVIGA-----HFQ 330
DB 455 KQDLQEAR---EARRAKOKLEIAIKTPYPMPNPIPPPLPDIPSFDIIADSLSDFK 510
QY 331 PTDVNVVHM-----GNVTEVR-----LPAFHSAGNDPPE 359
DB 511 DTDMKRLSMIEIEKXYEYMEKSKHLDQNLNETLTELKALKRKERTALDVLHSESS---D 567
QY 360 RHPVAVADRISDPQOKAR 378
DB 568 RGGPSSKHDTIKKPKQAGR 586
  
```

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:42:14 ; Search time 32.83 Seconds  
(without alignments)  
1273.191 million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165

Sequence: 1 MKPKIQRPDNEQAVARIG.....VLQYLFSPLRDRLRTTMRRE 435

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984.5	91.7	436	2 E87374	RsaA secretion sys
2	589	27.2	474	2 F98310	rhizobioicin secret
3	584	27.0	437	2 AE2972	secretion protein,
4	583	26.9	439	2 C96003	protein secretion
5	568	26.2	434	2 AB3099	hypothetical prote
6	564	26.1	436	2 G98187	rhizobioicin secret
7	564	26.1	443	2 B83221	metalloprotease
8	543.5	25.1	473	2 C95953	hypothetical prote
9	542.5	25.1	437	2 H97683	pse protein (Y127
10	542.5	25.1	437	2 A12908	HLyD family secret
11	530.5	24.5	511	2 D97753	alkaline protease
12	519.5	24.0	511	2 D71687	alkaline proteinas
13	514.5	23.8	432	2 S26697	apre protein - pse
14	514.5	23.8	432	2 G83489	alkaline proteinas
15	506.5	23.4	448	2 S12526	metalloprotease
16	476.5	22.0	442	2 AE0477	HLyD family secret
17	471.5	21.8	437	2 AB4933	hasa export system
18	438.5	20.3	452	2 AB2965	secretion protein,
19	438.5	20.3	497	2 C98318	alkaline proteinas
20	368.5	17.0	387	2 AG0835	probable type I se
21	368	17.0	481	2 D82381	secretion protein,
22	367	17.0	395	2 D83412	probable secretion
23	329	15.2	475	2 B81047	secretion protein,
24	325	15.0	391	2 D85547	probable membrane
25	325	15.0	391	2 H90686	probable membrane
26	314	14.5	473	2 F82561	hemolysin secretio
27	308.5	14.2	477	2 S48045	toxin apxIII secre
28	288.5	13.3	478	1 LEEDC	hemolysin export s
29	288	13.3	507	2 AD2140	hypothetical prote

30	267.5	12.4	477	2 B61378	leukotoxin secreti
31	267.5	12.4	478	2 S10058	hemolysin secretio
32	266	12.3	440	1 BYBRCD	cyad protein - Bor
33	264	12.2	438	2 T03513	probable secretion
34	262	12.1	478	2 E43599	toxin apxIII secret
35	261.5	12.1	454	2 AB1828	probable periplasm
36	260	12.0	500	2 AB2517	hemolysin secretio
37	251.5	11.6	479	2 T00229	hemolysin secretio
38	251	11.6	467	2 G82198	RTX toxin transpor
39	248	11.5	498	2 AH2468	hypothetical prote
40	234	10.8	478	2 D30169	leukotoxin secreti
41	215.5	10.0	583	2 S75805	hemolysin secretio
42	203.5	9.4	425	2 AB1018	probable type-I se
43	201.5	9.3	425	2 T14873	HLyD secretion pro
44	196	9.1	455	2 D95923	probable OMA famil
45	194	9.0	512	2 AB2047	hypothetical prote

ALIGNMENTS

Query Match	Best Local Similarity	91.7%	Score 1984.5;	DB 2;	Length 436;
Matches	408;	Conservative	3;	Mismatches	24;
				Indels	1;
				Gaps	1;
QY 1	1	1	1	1	1
QY 2	2	2	2	2	2
QY 3	3	3	3	3	3
QY 4	4	4	4	4	4
QY 5	5	5	5	5	5
QY 6	6	6	6	6	6
QY 7	7	7	7	7	7
QY 8	8	8	8	8	8
QY 9	9	9	9	9	9
QY 10	10	10	10	10	10
QY 11	11	11	11	11	11
QY 12	12	12	12	12	12
QY 13	13	13	13	13	13
QY 14	14	14	14	14	14
QY 15	15	15	15	15	15
QY 16	16	16	16	16	16
QY 17	17	17	17	17	17
QY 18	18	18	18	18	18
QY 19	19	19	19	19	19
QY 20	20	20	20	20	20
QY 21	21	21	21	21	21
QY 22	22	22	22	22	22
QY 23	23	23	23	23	23
QY 24	24	24	24	24	24
QY 25	25	25	25	25	25
QY 26	26	26	26	26	26
QY 27	27	27	27	27	27
QY 28	28	28	28	28	28
QY 29	29	29	29	29	29
QY 30	30	30	30	30	30
QY 31	31	31	31	31	31
QY 32	32	32	32	32	32
QY 33	33	33	33	33	33
QY 34	34	34	34	34	34
QY 35	35	35	35	35	35
QY 36	36	36	36	36	36
QY 37	37	37	37	37	37
QY 38	38	38	38	38	38
QY 39	39	39	39	39	39
QY 40	40	40	40	40	40
QY 41	41	41	41	41	41
QY 42	42	42	42	42	42
QY 43	43	43	43	43	43
QY 44	44	44	44	44	44
QY 45	45	45	45	45	45

```

QY      420 LFSPRLDTRLRTMREE 435
         |||
Db      421 LFSPRLDTRLRTMREE 436

```

RESULT 2  
F98310  
rhizoblocin secretion protein rspe (AF141932) [imported] - Agrobacterium tumefaciens (str. C58)  
C:Species: Agrobacterium tumefaciens  
C:Date: 22-Oct-2001 #sequence-Revision 22-Oct-2001 #text-change 11-Jan-2002  
C:Accession: F98310  
R:Goodner, B.; Hinkley, G.; Gattung, S.; Miller, N.; Blanchard, M.; Girello, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A:Reference number: A87359, PMID:11743194  
A:Accession: F98310  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-474 <KUP>  
A:Cross-references: GB:AA007870, PIDN:AAK90008.1, PID:915159979, GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR\_L\_2881  
A:Map position: linear chromosome

Query Match	27.2%;	Score 589;	DB 2;	Length 474;
Best Local Similarity	31.9%;	Pred. NO. 3.1e-29;		
Matches 140;	Conservative 85;	Mismatches 210;	Indels 4;	Gaps 2;

[illegible]

```

RESULT      3
AE2972
secretion protein, HlyD family [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C.Species: Agrobacterium tumefaciens
C.Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C.Accession: AE2972
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
erge, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCellan
; Karp, P.; Romero, P.; Zhang, S.

```

Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AE2972

A:Cross-References: GB:AE008669; PIDN:AAU4195.1; PID:q17741773; GSPDB:GN00187  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: prsE  
A:Map position: linear chromosome

Query Match	27.0%	Score 584	DB 2	Length 437
Best Local Similarity	31.9%	Pred. No. 5.7e-29		
Matches 139; Conservative	84	Mismatches 209	Indels 4	Gaps 2

QY	60	HEGGMIAKILVREEEKKAAQVLEIFDRPQANNAAGITRNOYALKAAMEARLLAERDOR	119
Db	62	HRKGQVEEIIIVNEBERKADYVMRIIDAITOTRNLALVTKRIDELTAARMARLEABDDL	121
QY	120	PEISFPADITSORADPMVYARAIADQAOFTERRTOTOGOVILMAAORLOYOSEETEGIDRO	179
Db	122	ALIFIPQALIAARDDPDVASAMRSTKLFERRKVSREGKKAQLEARITOFENHETIGLKAQ	181
QY	180	TOGLDKQGLFIDELIDLRKLYDKGLVPRPRLALLEARAGLSSTIGRLADRKAQVOGA	239
Db	182	EVAUYNGLAVALTEAEITTSOKSLREGQVSVQKRLNSIQTOAAFTGGEREKETIAYQAQTAGRI	241
QY	240	SPTOLKVRQIKQEFEEYOSSTETTERVALVTEKEVYASQAOKRIKIVSPVNTAONLR	289
Db	242	TETKIIOIILIDDELTEYGERELREIQAOMGEVERKKAABELKRIIDIVAPQOSMAYEMA	301
QY	300	PTEGASVYRAAPRVLVDIARPEDEAFVIOAHFOPTDVYVMGMVTEVRLRPAFHSAGNDPE	355
Db	302	VHTVGGVYVPRADPRIMLIVPDGDELEVOYIVPKRIDLOVQOKAMLRMTAENQVETPELE	361
QY	360	RHDPAAVADRISDPQOKARLFIQIVRVYVKQRLPYLRGRVYAAQMAQOIVPTGERTVIOY	419
Db	362	GHSRIAAADITTDQRLGSYLVARLISVSEREKLINAPLVGMPAAEAFIOTSERTALST	421
QY	420	LESPLDRLRTTMRER	435
Db	422	IAKPLTDJISRFAFREE	437

RESULT 4  
 C:Species: *Shinorhizobium meliloti*  
 C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
 C:Accession: C96003  
 R:Pinan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Herr  
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A:Title: The complete sequence of the 1,663-kb pSymb megaplasmid from the N2-fixing e  
 A:Reference number: A95842; MUID:21396508; PMID:11461431  
 A:Accession: C96003  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-439 <KUR>  
 A:Cross-references: GB:A591985; PIDN:CAC49691.1; PID:q1541178; GSPDB:GN00167  
 A:Experimental source: strain 1021, megaplasmid pSymb  
 R:Callbert, F.; Pinan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubba  
 pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001







C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: H97683  
R:Goodner, B.; Hinkle, G.; Gatlung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz, B. Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A:Reference number: A97359; PMID:11743194  
A:Accession: H97683  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-437 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK88425.1; PID:915157919; GSFDB:GN00169  
A:Gene: AGR\_C\_4904  
A:Map position: circular chromosome

Query Match 25.1%; Score 542.5; DB 2; Length 437;  
Best Local Similarity 30.8%; Pred. No. 2.1e-26;  
Matches 130; Conservative 83; Mismatches 204; Indels 5; Gaps 2;

OY 15 AAVAGCYIILTFGLGMAAFAPLDSAVTANGVSAE-VSDQYOHLEGGGLAKTLVRE 73  
DB 20 AVVVGLGLL---VGMGMAAFAPLDSAVTANGVSAE-VSDQYOHLEGGGLAKTLVRE 75  
OY 74 GEKVAAGVLELDPDPTQANAAAGITRNQYVALKAMEARLLAERDQPSISFPADLTQORA 133  
DB 76 GDRVAAGGILLRLSATVQANSLIENLTALQLYSRARLRMEAEPEFTYEDLTALTS 135  
OY 134 DPMVARAIADQAOFTERRQTIQGOVDLMAQRLOYSSEIIGIDRQGLKQDLQFIDE 193  
DB 136 SKSAFTFSDSEONFNSHRNALIGKKKOLATRKIQLADAEARGLDQVATENELAIKED 195  
OY 194 LIDRLKLDKGVPRPRLLALEARAGSLSGISGRITLDRSKAVQASDQTLKVRQKEF 253  
DB 196 VSKDELKKGGLVTLQRLNLKRLQSLNLEGGQGYTARAQTVGKLSLQDLQLDEDR 255  
OY 254 FEVQSQTITETRVRLAEATEKEEVVASDAQRIKITYSPANGTQONLEFTEGAVVAAEPL 313  
DB 256 KSEVTKDLTSEATVAEYEEERLAATRDQLDRIDRSPISAGRTYQLSVNINQVIGEV 315  
OY 314 VDIAPEDAEFYIQAHFOPTVDNVMGMWTEYRLPAFHSAGNPDEERHDPAAVADRISDP 373  
DB 316 MLVVPDKDLDLAEANITPRDIDQIVYGPVTRFPAFMQSTPDLSEAVAVAPDLQDS 375  
OY 374 OKQARLFGITRVNDKQPLHREGVTAGMPAQVIYVPGERGIVLYQLFSLPDLTKTMR 433  
DB 376 RTGTSYVLRIRPNKAGMGLPGKLYGMPAEVFIQTSERSVLSTYFVKPRDRLKTFV 435  
OY 434 EE 435  
DB 436 QE 437

RESULT 10  
AI2908  
HYD family secretion protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AI2908  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AI2908  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-437 <KUR>

A:Cross-references: GB:A008688; PIDD:AA143687.1; PID:g17741213; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: prfE  
A:Map position: circular chromosome

```
Query Match      25.1%, Score 542.5; DB 2; Length 437;  
Best Local Similarity 30.8%; Pred. No. 2.le-26;  
Matches 130; Conservative 83; Mismatches 204; Indels 5; Gaps 2;
```

OY 15 AAVRGVGIATLFEVLGLGMAFAFLDSAVIANGVSAE-VSDQVHLEGGMLKILIRE 73  
|| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db 20 AAIVVIGLGILL---VGWGMGAFAFLAGAVVATGVNVEGNSKIQTSLSGIVSEINVE 75  
  
OY 74 GEKVAGCVLFELDPDPTQANNAAGITRNOVALKAMEARLLAEORDRPSISFPADITSORA 133  
|::| |::| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db 76 GDRVAAGQILLRLSTTVQVANSTIENTLAQLSKRAALRAIEAPESPFTYEDLTALTLS 135  
  
OY 134 DPMVARIAIDEOAQETERRQTIGQVDLMNRQLQYSEIEGIDRQTGLDKDLGFIDE 193  
| | | | | | | | | | | | | | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db 136 SKSAKTFIDSEQNLPNSRRNALIGKKQLATRFKLDTADEARCLDVQVATEMELAIYKED 195  
  
OY 134 LIDLKLYDKGLVPFRPLLAEARAGSLSGTSGRLTADRKAAYOGASDTOLKVRIOKEF 253  
: : : | :| | :| | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db 196 VSKTELLBKKGVLTLQRNLNLKROLSNLEGGQGVIARAQTVCKLSLDLQLLDLEDGR 255  
  
OY 254 FEQVQSISTETFRVLAEYTEKEVNVSADQKRKITIYPNGNAONLRFTEGAVVRAAPEL 313  
: : : | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db 256 KSEVTIKDLSTIATVAYAEIEERLAATRDQLRIDLRSPLAGRIYOLSVININGVLOPGEVL 315  
  
OY 314 VDIADEDAEFVIOAHFOPTDYDNVMHGMVTEYRLPAFSAGNPDPERHDPPYAVALDRISDP 373  
| | | | | | | | | | | | | | | | | | | | | | :| :| :| :| :| :| :| :| :| :| :|  
Db 316 MLVVPDKODLMEANTPRDIQIVGQPVYRFAFNQSTTPDLASAVAVAAPDLQTDSS 375  
  
OY 374 QKOARLFLGIVRVNDKQLPHLRGRVTAAGMPAOQIVLPGERTVLOYLESPLRDILTMTMR 433  
: : : | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db 376 RGTGSTYYVLIRIPNKAGMGHLPGRKLYGPMPAEVFLOTSERSVLSYEVKPFQDDRILKKTFFV 435  
  
OY 434 EE 435  
:|  
Db 436 QE 437

RESULT 11  
D97753  
alkaline protease secretion protein AprE [imported] - Rickettsia conorii (strain Malil)  
C:Species: Rickettsia conorii  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001  
C:Accession: D97753  
R:Organa, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.: Science 293, 2093-2098, 2001  
A>Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
A:Reference number: A97700; WUID:21442074; PMID:11557893  
A:Accession: D97753  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-511 <KUP>  
A:Cross-references: GB:A006914; PIDD:AA102966.1; PID:g15619497; GSPDB:GN00173  
C:Genetics:  
A:Gene: aprE

Query Match 24.5%, Score 530.5; DB 2; Length 511;  
Best Local Similarity 30.0%; Pred. No. 1.5e-25;  
Matches 135; Conservative 88; Mismatches 202; Indels 25; Gaps 7;

OY 5 KIORTPDNQAVAR--IGYGIALLFEVGLG--WAAPFLDSAVIANGVSAEYS-QDVQ 59  
| : : | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db 68 KTDRKNNVQAARSPILFGIYVMTFLVIGLSALAPLDSGAVAVGVIMPSTNNKTIQ 127  
| : : | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|  
OY 60 HLEGMLAKILVREGKVKAGOVLELDPDQANNAAGITRNOVALKAMEARLLAEIRDOR 119

Db 128 HHGGIINAIYVKGQVKEGDKLIELETRIKSEHENILGOYRNFPLATENRLAERDNL 187  
QY 120 PSTSPADLTSSORADPVAARAIADEQAQFTERROTIGQVLDMAAQLQYQSEIEGIDRQ 179  
Db 188 EQLEFSQFLMODINLPVAKIITHTQENLFRSKEVYSEKDALHONIAQLQLEKKEGLEAK 247  
QY 180 TQGLKDLQGLFIEDELIDRLKYDKGLVPRRLALAEARAGSLSGSIGRLTADRKAQVGA 239  
Db 248 KVAASKTAEVYQDRKALRTLKEGQVQKAALLDQEAQKVAASKSDVATTEAETIGIRHAI 307  
QY 240 SDPQAKRQIKQEFFEVSQISITETRYRLAEVTEKEVVAADQAKRITVSPVNGTAQNLK 299  
Db 308 TETQIKINQNNYTERLTLELREAOVOTASLEKYNALDLSLNRVITIRAPVDGIVNSLK 367  
QY 300 FTEGAVVAAEPLVDIAPDEAFVIOAHFQPTDVNVHMGVTEVRLPAFHSAQNP--- 356  
Db 368 YHTIGVISHQPIRMEISPLINDPLIFAKYSQKNIDSVHGLVAKIRFSAFKSTTTFTT 427  
QY 357 -DEPRHDPVAVADRISDP-QKQARLFLGIVRD-----VKQLRPHLRGRVTAQMPA 405  
Db 428 GKAVYSIDPIVODERQYQGQODNYVAVRVEIDMDEFNKVAKYKNELEH-----PGMQA 481  
QY 406 QVTPGERTVLOYLFSPLDLTLTMTREE 435  
Db 482 EVOIVGTGRTLRLLDLPVDTAFKAFREK 511

## RESULT 12

D71687  
alkaline proteinase secretion protein apre (apre) RP314 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii  
C>Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
R:Accession: S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark, U  
Nature 396, 133-140, 1998  
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A:Reference number: A71630; MVID:99039499  
A:Accession: D71687  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-511 <AND>  
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:93868717; PIDN:CAA14774.1; PID:9386087  
A:Experimental source: strain Madrid E  
C:Genetic:  
A:Gene: apre; RP314

Query Match 24.08; Score 519.5; DB 2; Length 511;  
Best Local Similarity 29.3%; Pred. No. 7.1e-25;  
Matches 134; Conservative 93; Mismatches 192; Indels 39; Gaps 10;

QY 6 IORPTD---NFOAVAR--IGYGIALLFVGLLG--WAAFAPLDSAVIANGVASAEVS-Q 56  
Db 65 ITKTTDNRNNVAAQAASPIIFGIYVIFLVGLGWSALAPLDSGVAIGIVPSTNRK 124  
QY 57 DVQHLBEGMLAKILVREGKVKAGQVLELDPQTQANAAAGITRNOYVALKAMEARLAE 116  
Db 125 TIQHNIEGIIINAIYVKGDKYKDEKDLIELETRIKSEHENILSQYNNPLATENRLAE 184  
QY 117 DQRPISFPADLTSSORAD-PMAVARAIADQQAQFTERROTIGQVLDMAAQLQYQSEIEG 175  
Db 185 DNLEQIKF-SDFLMQNNINLEPAKIIHTQENLFRSKREVEVSEKDALHONIAQLQLEK 243  
QY 176 IDROTQGLKQGLFIEDELIDRLKYDKGLVPRRLALAEARAGSLSGSIGRLTADRNSKA 235  
Db 244 LEAKKIAASKTSEVYQDRKALRTLKEGQVQKAALLDQEAQKVAASKSDVATTEAETIAG 303  
QY 236 VQASDPTQLKRAQIKQEFFEVSQISITETRYRLAEVTEKEVVAADQAKRITVSPVNGTA 295  
Db 304 RHATTEQIKINQNNYTERLTLELREAOQVOTASLEKYNALDLSLNRVITIRAPVDGIV 363  
QY 296 QNLEFTEGAVVAAEPLVDIAPDEAFVIOAHFQPTDVNVHMGVTEVRLPAFHSAQNP 355

Db 364 NMLKYHTIGVISHQPIRMEISPLINDPLIEARISQKNIDSVHGLVAKIRFSAFKSR 423  
QY 356 P-----DPRHDPVAVADRISDPQKQARLFLGIVRD-----VKQLRPHLRG 397  
Db 424 PTTGKAVYSISP-----IVDERQHILGQQODNYVAVRVEIDMDEFNKVAKYKNELEH--- 476  
QY 398 RVTAGMPAOVIVPTGERTVLOYLFSPLDLTLTMTREE 435  
Db 477 ---PGMQAEVQIVTGTTRTLRLYLDLPVDTAFKAFREK 511

## RESULT 13

S26697  
apre protein - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 31-Mar-2000  
R:Accession: S26697  
R:Duong, F.; Lazdunski, A.; Cami, B.; Murgier, M.  
Gene 121, 47-54, 1992  
A:Title: Sequence of a cluster of genes controlling synthesis and secretion of alkali  
A:Reference number: S26696; MVID:93051361  
A:Accession: S26697  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-432 <DO>  
A:Cross-references: EMBL:X64558; NID:945279; PIDN:CAA45856.1; PID:945281  
C:Superfamily: hemolysin secretion protein D; lipoyl/biotin-binding homology  
C:Keywords: transmembrane protein  
F:48-91,284-313/Domain: lipoyl/biotin-binding homology #status atypical <LPB>

Query Match 23.8%; Score 514.5; DB 2; Length 432;  
Best Local Similarity 31.6%; Pred. No. 1.2e-24;  
Matches 137; Conservative 88; Mismatches 182; Indels 27; Gaps 11;

QY 15 AVARIGGIALLFVGLGMAAFAPLDS--AVIANGVASAEVSQDVQHLBEGMLAKILVR 72  
Db 11 AYARLGMILVLEFGGALMAAFAPLDQGVAVPATVILISQ-RKSYOHPLGQVKKHILVR 69  
QY 73 EGEKVKAGQVLELDPQTQANAAAGITRNOYVALKAMEARLAEHDDQRPISFPADLTSSOR 132  
Db 70 DQGVHEGEPLIRMEIPYQARANDSLNRYANANLQARQAEYDGRRTLEMPAGLAEQA 129  
QY 133 ADPMAVARAIADQQAQFTERROT-IOGVLDMAAQLQYQSEIEGIDROTQ---LKDQLG 188  
Db 130 PLPTLGBRL-ELQRLHSQRTALANELSLRANIEGRLQLEGL-RQTGNGNQLQRL- 186  
QY 189 FIEDELIDRLKYDKGLVPRRLALE-----ARAGSLSGSIGRLTADRKAQVQASD 241  
Db 187 -LNSQLSGARDLAEQGYMPRQMLEOERQLAEVNARLSESSGRFGQIR-----QSI 238  
QY 242 TQLKVRQIKQEFFEVSQISITETRYRLAEVTEKEVVAADQAKRITVSPVNGTAQNLK 301  
Db 239 AQMRIAQREERYKREKVGQAEIQTQVNAKRTLMEELSSARTELIRNAETRAPVSGYAGIKVF 298  
QY 302 TEGAVVAAEPLVDIAPDEAFVIOAHFQPTDVNVHMGVTEVRLPAFHSAQNPDEPRH 361  
Db 299 TDGVAVIGPGLMLMTIVNSDSLEVEGQAVNLVDRINSGLRVEVLELFTAFNQSRTPV 358  
QY 362 DPVAVADRISDPQKQARLFLGIVRVYKQLRPHLRG-RVTAGMPAOVIVPTGERTVLOY 420  
Db 359 VTMVASADRLDEQNKQKQYVYRVAQVDAAM-GKLKGQIRPGMAVQVYFRTGERSLNYL 417  
QY 421 FSPLDLTLTMTREE 434  
Db 418 FKPLFDRHVALAE 431

## RESULT 14

G83489  
alkaline proteinase secretion protein Apre PA1247 [imported] - Pseudomonas aeruginosa  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:49:04 ; Search time 17.51 Seconds

(without alignments)  
961.908 million cell updates/sec

Title: US-09-913-414-5

Perfect score: 2165  
Sequence: 1 MKPKIQRPDTNFOAVARIG.....VLQYLSPDRDLFTTMRRE 435

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	514.5	23.8	432	1	APRE_PSEAE 003025 pseudomonas
2	506.5	23.4	448	1	PRTB_ERWCH P23537 erwania chr
3	313.5	14.5	477	1	RT3D_ACTPL 008633 actinobacil
4	288.5	13.3	478	1	HLV4_ECOLI 009986 escherichia
5	267.5	12.4	477	1	HLVD_ACTAC P18790 actinobacil
6	267.5	12.4	478	1	HLVD_ECOLI P06739 escherichia
7	266	12.3	440	1	CTAD_BORPE P11091 bordetella
8	243.5	11.2	478	1	HLVD_PASBP P55125 pasteurella
9	233.5	10.8	477	1	RT1D_ACTPL P26761 actinobacil
10	233	10.8	478	1	HLVD_PASHA P16534 pasteurella
11	164	7.6	424	1	MCHE_ECOLI 09exn6 escherichia
12	163	7.5	413	1	CVAA_ECOLI P22519 escherichia
13	147.5	6.8	343	1	YUCH_ECOLI P32716 escherichia
14	146	6.7	421	1	EXOF_RHIME 002718 rhizobium m
15	142	6.6	352	1	YH1I_ECOLI P37626 escherichia
16	136.5	6.3	385	1	Y894_HAEIN 057500 haemophilus
17	130.5	6.0	879	1	CTPI_MYCTU 002171 onchocerca
18	130	6.0	1625	1	MYSP_ONCYO 010900 mycobacteri
19	129.5	6.0	848	1	MYSP_DIRIM P13382 dirofilaria
20	128	5.9	371	1	MACA_ECOLI P75850 escherichia
21	128	5.9	882	1	MYSP_CAEEL P10570 caenorhabdi
22	126	5.8	390	1	EMRA_ECOLI P27303 escherichia
23	124	5.7	371	1	MACA_YERPE P58411 yersinia pe
24	124	5.7	880	1	MYSP_BRUMA 001202 brugia mala
25	123.5	5.7	1576	1	RSPC_AOUPY 09x652 aquifex pyt
26	122.5	5.7	378	1	YIBH_ECOLI P32107 escherichia
27	122	5.6	371	1	MACA_ECO57 P58410 escherichia
28	121.5	5.6	378	1	YIAV_ECOLI P37683 escherichia
29	121.5	5.6	545	1	HTRS_HAUNI 048318 halobacteri
30	121.5	5.6	4684	1	PIEL_HUMAN 015149 homo sapien
31	120	5.5	572	1	YOG2_DEIRA 09rm6 delinococcus
32	120	5.5	692	1	MTS_PODCA 005000 podocoryne
33	120	5.5	4473	1	PLEI_CRIGR 09j155 cricetulus

34	119.5	5.5	474	1	LCND_LACIA 000565 lactococcus
35	118.5	5.5	440	1	ODP2_ZYMO 066119 zymomonas m
36	118	5.5	863	1	MYSP_TAESO P35418 taenia soli
37	117.5	5.4	1375	1	RPOB_VIBCH 09kv30 vibrio chol
38	117	5.4	652	1	RPSD_CAVCR P52324 caulobacter
39	116.5	5.4	449	1	COMB_STRPN P36458 streptococ
40	116	5.4	1238	1	SBCC_RHOCA 068032 rhodobacter
41	115	5.3	398	1	ACOC_BACSU 031550 bacillus su
42	115	5.3	995	1	HIPI_HUMAN 000291 homo sapien
43	114.5	5.3	387	1	EMRK_ECOLI P52599 escherichia
44	114.5	5.3	390	1	EMRA_HAEIN P44928 haemophilus
45	114.5	5.3	412	1	MTRC_NEIGO P43505 neisseria g

## ALIGNMENTS

```
RESULT 1
APRE_PSEAE STANDARD: PRT; 432 AA.
ID APRE_PSEAE
AC 003025:
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alkaline protease secretion protein apre.
GN APRE OR PA1247.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
OC [1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 15692 / PA01;
RC MEDLINE=93051361; PubMed=1427098;
RX MEDLINE=93051361; PubMed=1427098;
RA Dong F., Lazdunski A., Cami B., Murgier M.;
RT "Sequence of a cluster of genes controlling synthesis and secretion
RT of alkaline protease in Pseudomonas aeruginosa: relationships to
RT other secretory pathways.";
RL Gene 121:47-54(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stever C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.V., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC - FUNCTION: INVOLVED IN THE SECRETION OF ALKALINE PROTEASE.
CC - SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL).
CC - SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

EMBL: X64588; CNA45856.1; -  
DR EMBL: AE004554; AAG04636.1; -  
DR PIR: S26697; S26697.  
DR HSSP: P02905; IBD0.  
DR InterPro: IPR002215; HLVD.  
DR InterPro: IPR003997; RtxD.  
DR Pfam: PF00529; HLVD; 1.  
DR PRINTS: PR01490; RXTXOIND.  
DR PROSITE: PS00543; HLVD\_FAMILY; 1.



[illegible]

```

Db 100 KPIENSIIVEIKESVARKGVLLKL--TALGAEDTLTKQSSLSLQARLEQIRQIILS 157
OY 115 ---ERDQPSISPADLTQSQR-ADPMVARAIADQAF--TERQTIQGYDL--MNAQRL 167
Db 158 RSTELMKLPFLKLPDEEYKPNVSEEEVRLSLIKEDFSTWQMOQKOKELNLDKKRAERL 217
OY 168 QYSEIEGIDRQGTQKLDQGLFIEDELIDRLKLYDKGLVPRPLALEAR---AGSLSG 223
Db 218 TILAIRIRYENVSREKSRDL-----DPSRLIKQAIKHAHVLQEDENKYEVAANLRY 270
OY 224 SIQRILADREKAVQASDPLQKVRQIKQEFEEQVSQITFRVRLAVTKEVEVASDAQK 283
Db 271 YKSOLEIESEILSAKEEYQLVTLQFKNELDLKQRTTDSIELTLEKNE-----ERQ 326
OY 284 RIKIVSPVNGTAQNLRFETEGAVVRAEPLVDIAPDEDAVIAQHPOPTDVNVMGMVT 343
Db 327 ASVIRAPVSKYQGLKHTGGVVTATETLMVIVPEDDTLVTLQNKDIGFVNGQNA 386
OY 344 EVRLPAFHSAGNPDPERHADVANA-----DRISDPQKARLFLGIVRDYKQLP---P 393
Db 387 IIRVEAF-----PYTRYGYLVGKVKKNIMLDAIED--QKGLVFENVIVSYEENDLSTGNK 438
OY 394 HLRGRVAGMPAQVIPTGERTVQLYLFSPLRDLRTTMBE 434
Db 439 HI--PLSSGMAVTAETIKTGMRSVISTLSPLESVTESLHE 477

RESULT 5
HLVD_ACTAC STANDARD; PRT; 477 AA.
AC P18790;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, last sequence update)
DT 01-OCT-1996 (Rel. 34, last annotation update)
DE Leukotoxin secretion protein D.
GN LKTD OR AALTD.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OC actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_Taxid=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JP2;
RX MEDLINE=90384843; Pubmed=2402458;
RA Guttmiller J.M., Kraig E., Cagle M.P., Kolodrubetz D.;
RT "Sequence of the lktD gene from Actinobacillus
RT actinomycetemcomitans."
RL Nucleic Acids Res. 18:5292-5292(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92072446; Pubmed=1961107;
RA Lally E.T., Golub E.E., Kleba I.R., Taichman N.S., Decker S.,
RA Berthold P., Gibson C.W., Demuth D.R., Rosenblom J.;
RT "Structure and function of the B and D genes of the Actinobacillus
RT actinomycetemcomitans leukotoxin complex."
RL Microb. Pathog. 11:111-121(1991).
CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF THE HEMOLYSIN/LEUKOTOXIN.
CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
CC -1- SIMILARITY: THE N-TERMINAL (AA 1-153) IS SIMILAR TO E.FAECALIS
CC PORE FORMING PROTEIN EBSA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X53956; CA37907.1; -
CC PIR: S11215; S11215.

```

```

DR PIR; B61378; B61378.
DR InterPro; IPR002215; HLVD.
DR InterPro; IPR003997; RtxD.
DR Pfam; PF00529; HLVD; 1.
DR PRINTS; PR01490; RTXTOXIND.
DR PROSITE; PS00543; HLVD_FAMILY; 1.
KW Hemolysin; Cytolysin; Transport; Transmembrane; Inner membrane.
FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 60 80 POTENTIAL.
FT DOMAIN 81 477 PERIPLASMIC (POTENTIAL).
SQ SPOUNCE 477 AA; 54651 MW; 54F20128CADE6260 CRC64;

Query Match 12.4%; Score 267.5; DB 1; Length 477;
Best Local Similarity 23.7%; Pred. No. 5.9e-10;
Matches 108; Conservative 87; Mismatches 207; Indels 53; Gaps 15;

OY 6 IORPTDFOAVARIGVGIATLFGVGLGMAAPPLDPAVITANGVS-AEVSODVOMLEGG 64
Db 49 IETPVSN--APREVSIMLFLTALV--VSTFSVEIATASGKFKALSGRSKEIKPIENS 105
OY 65 MLAKIIVREGEKVKACGVLEFLDPTQANAAAGITRNQVALKAMEARL-----LAERDQR 119
Db 106 LVKHIFVKEGEYKKGGLLKLALGAEADTLTKTSLQAKLEEFYKSLLEAVERDQL 165
OY 120 PSISF-PADLTQSADPMVARAIADQAF--TERQTIQGYDL--MNAQRLQYSEIEGI 176
Db 166 PILDFSKIDLPFTMNQKRVTLIEQFSTWQKQHQKTLNKKRAEKLSTYARL--- 222
OY 177 DRQYQGLKDLQGFTEDE---LIDRLKLYDKGLVPRPLALEARAGSLSSIGRLTADRS 233
Db 223 -----KRYEGLINTQVRLDDEFRALYKEHAIAKHIVLDEE---NKYQDAINLEYYKA 272
OY 234 KAVQ-----GASDTQLKVRQIKQEFEEQVSQITFRVRLAEVTEKEVEVASDAQRIK 286
Db 273 SIMQVENEVLAKEEQELVQGLFKNDILDKLQKQTDVNNILTFELDKN---NQRQGVSE 328
OY 287 IVSPVNGTAQNLRFETEGAVVRAEPLVDIAPDEDAVIAQHPOPTDVNVMGMVTEVR 346
Db 329 IRAPVSGVQDLKVTHTIDGVVTTAETLMVAVPEDSLEVALIIONKKGIFVKEQOEYVIR 388
OY 347 LPAF-----HSAGNPDPERHADVANAADRISDPQKARLFLGIVRDYKQLPRLRG-R 399
Db 389 VEAFPPYTRYGYLTGKVKV-----ITDAILHP-KLGLVFVFTIELDKTLSTEEKELPL 441
OY 400 TAGMPAQVIPTGERTVQLYLFSPLRDLRTTMBE 434
Db 442 SAGMEITAEIKTGMRSVISTLSPLESIDKSLRE 476

RESULT 6
HLVD_ECOLI STANDARD; PRT; 478 AA.
AC P06739;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Hemolysin secretion protein D, plasmid.
GN HLVD.
OS Escherichia coli.
OC Plasmid Inci7 pHLV152.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Hess J., Wels W., Vogel M., Goebel W.;
RT "Nucleotide sequence of a plasmid-encoded hemolysin determinant and
RT its comparison with a corresponding chromosomal hemolysin sequence."
RL FEBS Microbiol. Lett. 34:1-11(1986).
RN [2]
RP TOPOLOGY
RX MEDLINE=92357011; Pubmed=1495479;

```





```

Db      376 VDDROGHSYRTIALAHAPLEVDGK---PRL---LKEGNAVQADIRTGSRILEYLLSPV 429

RESULT 8
ID      HLVD-PASSP      STANDARD:      PRT:      478 AA.
AC      P53125;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      Leukotoxin secretion protein D.
GN      LKTD.
OS      Pasteurella haemolytica-like sp. (strain 5943B).
OC      Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC      Pasteurella.
OX      NCBI_TaxID=28165;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93239320; PubMed=8478098;
RA      Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
RT      "Molecular characterization of a leukotoxin gene from a Pasteurella
RT      haemolytica-like organism, encoding a new member of the RTX toxin
RT      family.";
RL      Infect. Immun. 61:2089-2095(1993).
CC      -!- FUNCTION: INVOLVED IN THE TRANSPORT OF THE LEUKOTOXIN.
CC      -!- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL).
CC      -!- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L12148; AAA16446.1; -
DR      InterPro: IPR002215; HLVD.
DR      InterPro: IPR003997; RTX.D.
DR      Pfam: PF00529; HLVD; 1.
DR      PRINTS: PR01490; RTXTOXIND.
DR      PROSITE: PS00543; HLVD_FAMILY; 1.
KW      Hemolysis; Cytolysis; Transport; Transmembrane; Inner membrane.
FT      DOMAIN 1 77 CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM 78 98 POTENTIAL.
FT      DOMAIN 99 478 PERIPLASMIC (POTENTIAL).
SQ      SEQUENCE 478 AA; 54650 MW; 221461A69482913A CRC64;

Query Match      11.2%; Score 243.5; DB 1; Length 478;
Best Local Similarity 24.4%; Pred. No. 1.9e-08;
Matches 113; Conservative 73; Mismatches 174; Indels 103; Gaps 18;

OY      31 LLGNAAPPLDSDAVIANGVSAEV-----SQDVHLEGGLAKITLYREGKV 77
DB      59 LIAVLIMFLAVLAVLASVSEVEIYAPAGRLTFSGRSKETPTVQIEFVKGQGV 118
OY      78 KAGVLFELDP--TGANAAGITRNQVYALKAMEARLL--AENDORSISFPADLT--- 129
DB      119 EKGQLVSLTALGSDADIKTITTSIAKLENYRQITLTIEKESLPVY---DLSTTE 174
OY      130 ---SQRADPVYARAIDE-----QAOFTER-----ROTTQGVVDLMAAORLOYOS 171
DB      175 FKDSEEDRLIKHLEIQYITWQOKOTKLAVKRKAERQOTISA-----YVR 223
OY      172 EIEGIDROTGKIQDLGTFIEDELIDKLKLYKGLVPRRLALAEARAGSLSGIGRLTAD 231
DB      224 KYEGATRIEQ-----EKLDFKRLYQOKSLSKHEHLLSOENKAIEAQ---NELAVY 270
OY      232 RSKAVOGASD-----TQLKVRQIKOEFEQVOSISITETVRLAVEYKEKEVASDAOKR 284
DB      271 RSKLNEESDLINVKELLETITQFFKSDVLEKLAQHIENRQOLEL-EKN---NORROA 326

```

```

OY      285 IKIVSPVNGTAQNLKRFTEGAVYRAAEPLVDIAPDEAEVIOAHPOPTDVNVHMGVTE 344
DB      327 SMIRAPVSGTVOQLKTHITIGVTTAETLMTIPEDEVDLEATLQNKDIGEFAAGOEVI 386
OY      345 VRLPAPFASAGNDPERRHDPAP-----VADRIS-----DPKOARL---FLGIVRDVQL-- 391
DB      387 IKVETP-----PYTRYGYITGRKIKHSPALIEOPNGLGVFNATVSDIKQALSS 434
OY      392 PPHLRGRVTAGMPAQVIVPTGERTVLOYLEFSPRLDRLRTTMR 434
DB      435 PDGKHTELGLGWTITAEIKTGRSVMSYLLSPLEESVTESLRE 477

RESULT 9
ID      RTID-ACPL      STANDARD:      PRT:      477 AA.
AC      P26761;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      RTX-I toxin determinant D (APX-ID) (HLV-ID) (Cytolysin ID) (CLY-ID)
DE      (Toxin RTX-I secretion protein D).
GN      APXID OR CLYID OR HLVD OR APPD.
OS      Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC      Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC      Actinobacillus.
OX      NCBI_TaxID=715;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=SEROTYPE 5;
RX      MEDLINE=91317735; PubMed=1860823;
RA      Chang Y.-F., Young R., Struck D.K.;
RT      "The Actinobacillus pleuropneumoniae hemolysin determinant: unlinked
RT      appca and appbd loci flanked by pseudogenes.";
RL      J. Bacteriol. 173:5151-5158(1991).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      STRAIN=ISOLATE CVI 13261 / SEROTYPE 9;
RX      MEDLINE=92040145; PubMed=1937809;
RA      Smits M.A., Briatore J., Jansen R., Smith H.E., Kamp E.M.,
RA      Gielkens A.L.;
RT      "Cytolysins of Actinobacillus pleuropneumoniae serotype 9.";
RL      Infect. Immun. 59:4497-4504(1991).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      STRAIN=S 4074 / SEROTYPE 1;
RX      MEDLINE=94237497; PubMed=8181764;
RA      Frey J., Haldemann A., Nicolet J., Boffini A., Prentki P.;
RT      "Sequence analysis and transcription of the apxi operon (hemolysin I)
RT      from Actinobacillus pleuropneumoniae.";
RL      Gene 142:97-102(1994).
CC      -!- FUNCTION: INVOLVED IN THE TRANSPORT OF THE TOXIN RTX-I AS WELL AS
CC      THAT OF RTX-II.
CC      -!- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL).
CC      -!- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 5.
CC      -!- SIMILARITY: BELONGS TO THE HLVD FAMILY OF SECRETION PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: M65808; AAB00967.1; -
DR      EMBL: X61112; CAA43426.1; -
DR      EMBL: X68595; CAA48588.1; -
DR      PIR: B40366; B40366.
DR      PIR: S18856; S18856.
DR      InterPro: IPR002215; HLVD.
DR      InterPro: IPR003997; RTX.D.
DR      Pfam: PF00529; HLVD; 1.

```



```

OY      289 SPVNTAQNLFRETCGAVRAAEPLVDIAEDAEVAIDQHPRTVDNVHMGMTVERLP   348
Db      331 APVSCTVOOLKHTHTIGCVYTAAETLMTIIVPEDDVEAFALPVPNKDIGFAAGQEVIITKE   390
OY      349 AF-----HSAG-----NDPERHDPAVNADRISDPQAKRFLGLIVRDVKQL--PHLL   395
Db      391 TEPYRYGVLTGRIKHIKISPDALEQPNVG-----VFNNFTIADRNKLNSPPGR   438
OY      396 RGRVTAGMPAOYIVPTGERTVLOLYFSPRLDTLRITMKHE   434
Db      439 KIDLSSGMTTIAEIKTGERSVMSYLLSPLEESVTESLSE   477

RESULT 11
ID      MCHB_ECOLI          STANDARD;             PRT;       424 AA.
AC      O9EXN6;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DM      01-MAR-2002 (Rel. 41, Last annotation update)
DE      Microcin H47 secretion protein mchE.
GN      MCHB.
OS      Escherichia coli.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC      Escherichia.
OX      NCBI_TaxID=562;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-H47;
RX      MEDLINE=21091907; PubMed=11181394;
RA      Azpiroz M.F., Rodriguez E., Lavina M.;
RT      "The structure, function, and origin of the microcin H47 ATP-binding
RL      cassette exporter indicate its relatedness to that of colicin V.";
CC      Antimicrob. Agents Chemother. 45:969-972(2001).
CC      - FUNCTION: PROBABLY INVOLVED, IN CONJUNCTION WITH MCF, IN THE
CC      SECRETION OF MICROCIN H47
CC      - SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (BY SIMILARITY).
CC      - SIMILARITY: BELONGS TO THE HLTD FAMILY OF SECRETION PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      The European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AJ278866; CAC21493.1; -.
DR      InterPro; IPR002215; HLTD.
DR      Pfam; PF00529; HLTD; 1.
DR      ProSite; PS00543; HLTD_FAMILY; 1.
KW      Transport; Protein transport; Bacteriocin transport; Transmembrane;
KW      Inner membrane.
FT      DOMAIN              1         25        CYTOPLASMIC (POTENTIAL).
FT      TRANSMEM           26         46        POTENTIAL.
FT      DOMAIN            47         424        PERIPLASMIC (POTENTIAL).
SO      SEQUENCE           424 AA; 48719 MW; 3CAD0D7DEC3A41F4 CRC64;

Query Match               7.6%; Score 164; DB 1; Length 424;
Best Local Similarity     21.7%; Pred. No. 0.0017;
Matches    93; Conservative 73; Mismatches 142; Indels 120; Gaps 18;

OY      50 VSAEYS---ODVQHLEG--GMLAKTLVRGEKVKAAGVLFELDPQTQAANAAGITENQYA   104
Db      55 VSGETTYTPRAVNATISGVGGFVRRPFVHEGOALKKGDPYLIDISKSTSGVITDNH---   111
OY      105 LKAMEARLARERDORPSISFPADLTQS--RADPMVARAIADBOAQFTERRORTIOGOVDLM   162
Db      112 -----RRDIENQLRVYDNIISR--LESKKIT-----LDLTL   140
OY      163 NAQRLOY-----OSEIECIDROTQGLKDOLFIEDELIDLRKLYDKGIWPRRL---LA   213
Db      163 NAQRLOY-----OSEIECIDROTQGLKDOLFIEDELIDLRKLYDKGIWPRRL---LA   213

```

```

Db      141 EKQRLQYIDAFRRSSDI--IQRAEGIK-----IKNNMNEYRNVAQKLIKNDQLTNGVA 194
Qy      214 L----EARGSLSG-----SIGRLTRADSKKAVAGASDIOAKYRKQIKQEFFEEDVOSIETFR 265
Db      195 LYYOQNNMLLSGSCNEONALITLLEFSQIOQTQADAFENRIYOMELORYE-----244
Qy      266 VRLAEVTEKEVAVSAOKRIKIVSVNQTQGNLRFEEGAVVAARAEPVLDIAPDEAVI 325
Db      245 -----LQKELVNTDVEGELLITRALTDSKVDLS-VTVGQWNTGDSILOQIPENIENY 297
Qy      326 QAHEPPTD-VDNVHMGVTEVRLPAFHS-----AGNEDP 358
Db      298 LILWPNDAVPIISGDKVNIITYEAFRAFKGQFSATYKITSIRTPASQEMLYTKGAPQN 357
Qy      359 EKHDEVAVADRTSDFOKARFLGLIVRDVQKLPRLHGRVATAGMPAQVIVPGEFTVLO 418
Db      358 TPQASVPWYKVIAMPEKQ-----IIRYDEKYLTP-----LENGKAKESTFLLEKRIYQ 405
Qy      419 YLFSPLRD 426
Db      406 WMLSPFYD 413

RESULT 12
CVAA_ECOLI STANDARD: PRT: 413 AA.
AC P22519;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Colicin V secretion protein cvaa.
GN CVAA.
OS Escherichia coli.
OC Plasmid IncFI ColV3-K30.
CC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91065315; PubMed=2249654;
RA Gilson L., Mahanty H.K., Kotler R.;
RT "Genetic analysis of an MDR-like export system: the secretion of
RT colicin V.";
RL EMBO J. 9:3875-3884(1990).
RN [2]
RP TOPOLOGY.
RX MEDLINE=96042091; PubMed=7592380;
RA Skvirsky R.C., Reginald S., Shen X.;
RT "Topology analysis of the colicin V export protein cvaa in Escherichia
RT coli.";
RL J. Bacteriol. 177:6153-6159(1995).
CC -1- FUNCTION: INVOLVED, IN CONJUNCTION WITH CVAB, IN THE SECRETION OF
CC COLICIN V.
CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-BOUND (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE HLXD FAMILY OF SECRETION PROTEINS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X57524: CAA40743.1; -.
DR PIR: S12271; IRECSA.
DR InterPro: IPR002215; HLXD.
DR Pfam: PF00529; HLXD.1.
DR PROSITE: PS00543; HLXD_FAMILY.1.
KW Plasmid; Transport; Protein transport; Bacteriocin transport;
KW Transmembrane; Inner membrane; Alternative initiation.
FT CHAIN 1 413 COLICIN V SECRETION PROTEIN CVAA.
FT CHAIN 161 413 COLICIN V SECRETION PROTEIN CVAA*.

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 19, 2002, 10:48:41 ; Search time 53.06 Seconds  
(without alignments)  
1418.258 Million cell updates/sec

Title: US-09-913-414-5  
Perfect score: 2165  
Sequence: 1 MKPKIQPTDNEQAVARIG.....VLQYLFSPDLRLRTTMR 435

Scoring table:  
BIOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP:REMBL\_19:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*  
16: sp\_virus:\*  
17: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1984.5	91.7	436	16	O85351
2	1972.5	91.1	436	2	O9RM9
3	606	28.0	436	16	O92NX2
4	589	27.2	436	2	O9X6N7
5	583	26.9	439	16	O33679
6	564	26.1	443	16	O9HYJ9
7	560.5	25.9	434	16	O98IG9
8	543.5	25.1	473	16	P96438
9	540.5	25.0	443	2	O54417
10	536.5	24.8	473	2	O05198
11	530.5	24.5	511	16	O921J2
12	522	24.1	443	2	O9XB63
13	521.5	24.1	443	2	O54457
14	519.5	24.0	511	16	O92DL4
15	512	23.6	435	2	O05694
16	510	23.6	433	2	O92G93

17	509	23.5	437	2	O67994	O67994 pseudomonas
18	502.5	23.2	447	16	O988E5	O988E5 rhizobium 1
19	499	23.0	435	2	O9FCN8	O9FCN8 rhizobium 1
20	495	22.9	444	2	O9ZNT8	O9ZNT8 pseudomonas
21	493	22.8	437	2	O87809	O87809 pseudomonas
22	489.5	22.6	427	2	O921G7	O921G7 campylobact
23	482.5	22.3	427	2	O9R8D9	O9R8D9 campylobact
24	479.5	22.1	443	2	O9RHT1	O9RHT1 pseudomonas
25	479	22.1	439	2	O9R9H4	O9R9H4 pseudomonas
26	477	22.0	438	2	O9KCS5	O9KCS5 pseudomonas
27	476.5	22.0	450	16	O9ZLS8	O9ZLS8 rhizobium m
28	471.5	21.8	437	2	O57387	O57387 serratia ma
29	467.5	21.6	440	16	O98140	O98140 rhizobium 1
30	467.5	21.6	441	2	O85376	O85376 proteins mir
31	436	20.1	433	2	O52859	O52859 rhizobium 1
32	368	17.0	481	16	O9KRM3	O9KRM3 vibrio chol
33	367	17.0	335	16	O912M0	O912M0 pseudomonas
34	329	15.2	475	16	O9JY53	O9JY53 neisseria m
35	326.5	15.1	475	2	O9X580	O9X580 neisseria m
36	326.5	15.1	475	2	O9JPK4	O9JPK4 neisseria m
37	314	14.5	473	16	O9PAU7	O9PAU7 xyliella fas
38	288.5	13.3	464	2	P74828	P74828 sphingomona
39	283.5	13.1	439	2	O87505	O87505 escherichia
40	276.5	12.8	478	2	O08269	O08269 escherichia
41	272.5	12.6	468	16	O98MG5	O98MG5 rhizobium 1
42	271.5	12.5	477	2	O9RCG6	O9RCG6 pasteurella
43	264	12.2	438	2	O68080	O68080 rhodobacter
44	261.5	12.1	454	16	O9JTT34	O9JTT34 neisseria m
45	251.5	11.6	479	2	O82886	O82886 escherichia

#### ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	436 AA.
O85351	O85351			
AC	O85351			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	MEMBRANE FORMING UNIT (RSA SECRETION SYSTEM, MEMBRANE PROTEIN RSAB).			
GN	RSAB OR CCI1009.			
OS	Caulobacter crescentus.			
OC	Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;			
OX	NCBI_TaxID=69394;			
RN				
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 19089 / CH15, AND NA1000;			
RA	Avram P., Smit J.K.;			
RT	"The Caulobacter crescentus paracrystalline S-layer protein is secreted by an ABC transporter (type I) secretion apparatus.";			
RL	J. Bacteriol. 180:3062-3069(1998).			
RN				
[2]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 19089 / CH15, AND NA1000;			
RA	Avram P., Smit J.K.;			
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.			
RN				
[3]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 19089 / CH15, AND NA1000;			
RA	Avram P.;			
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.			
RN				
[4]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 19089 / CH15;			
RA	MEDLINE=21173698; PubMed=11259647;			
RX	Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E., Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,			



RA DeBooy R.T., Dodson R.J., Durrkin A.S., Gwinn M.L., Haft D.H.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Barry K.,  
RA Utechtack T., Tian K., Wolf A., Vamathevan J., Emolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,  
RT "Complete genome sequence of *Caulobacter crescentus*,"  
PL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
DR EMBL; AF062345; AAC38667.2; -  
DR EMBL; AE005779; AAC2293.1; -;  
DR TIGR; CCI009; -;  
DR InterPro: IPR002215; HlyD.  
DR InterPro: IPR003997; RtxD.  
DR Pfam: PF00529; HlyD; 1.  
DR PRINTS: PR01490; RXTXOIND.  
DR PROSITE: PS00543; HLYD\_FAMILY; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 436 AA; 48389 MW; 90E52A834673C9DC CRC64;

Query Match	91.7%;	Score 1984.5;	DB 16;	Length 436;
Best Local Similarity	93.6%;	Pred. No. 1.2e-117;		
Matches 408; Conservative	3;	Mismatches 24;	Indels 1;	Gaps 1;

Oy	1	MKPKRIORPMDNPOAVARIGGIIATLFEVGLGMAAFAPLDSAVINGVSAFVS -ODVQ	59
		:	
Dd	1	MKPPRIQPPIDNFOAVARIGGIIATLFEVGLGMAAFAPLDSAVIANGVSAEGNKRTYO	60
Oy	60	HLEGGMLAKLIVRGEKVKAGOVLFELDPTQAAAAGITRNOYVALKAMEFARLLAERDOR	119
Dd	61	HLEGGMLAKLIVRGEKVKAGOVLFELDPTQAAAAGITRNOYVALKAMEFARLLAERDOR	120
Oy	120	PSISPRADLTQSQRADPMYARAIADEQAQFIERRQCTOGQVUDLMAAQRIQYOSPIBESIDQO	179
Dd	121	PSISPRADLTQSQRADPMYARAIADEQAQFIERRQCTOGQVUDLMAAQRIQYOSPIBESIDQO	180
Oy	180	TQGLKQDQGFIEDELDLIRKLYDKGLVPRRLLALAEARGLSGSIGRLTADRSKAVQGA	239
Dd	181	TQGLKQDQGFIEDELDLIRKLYDKGLVPRRLLALAEARGLSGSIGRLTADRSKAVQGA	240
Oy	240	SDTQLTQKQIKOEFPFEQVOSISITETRVRRLAEVTEKEVYASDAQRIKITYSPVNGTQONLR	299
Dd	241	SDTQLTQKQIKOEFPFEQVOSISITETRVRRLAEVTEKEVYASDAQRIKITYSPVNGTQONLR	300
Oy	300	PFTFEGAAVRAAEPLVDIPDEDEAEFVIOAHFOPTVDVNVHMGMTYEVRLPAFHSAGNPDEE	359
Dd	301	PFTFEGAAVRAAEPLVDIPDEDEAEFVIOAHFOPTVDVNVHMGMTYEVRLPAFHSAREIPILN	360
Oy	360	RHDPVAVADRISSDPQOARLEFLGIVRVYDKOLPRHLRGVATAGMPAOVIVPTGERTVLOY	419
		:	
Dd	361	GTIQTSLSDRISDPSQNKLDYETFLGIVRVYDKOLPRHLRGVATAGMPAOVIVPTGERTVLOY	420
Oy	420	LFSPLADLTLTMTMREE 435	
Dd	421	LFSPLADLTLTMTMREE 436	

## RESULT 2

ID	Q9RM9	PRELIMINARY;	PRT;	436	AA.
----	-------	--------------	------	-----	-----

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-JUN-2001 (Tremblay, 17, Last annotation update)

GN RSAE.

0C Bacteria; proteobacteria; alpha subdivision; caulobacter group;  
0C caulobacter

PN	11
OX	NCRLTAXID=09394;

SEQUENCE FROM N.A.  
STRAIN=JS4000:  
RC

The secretion signal of the *Caulobacter crescentus* S-layer protein is

RT located within the C-terminal 82 amino acids of the molecule.  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF193064; AAF07963.1; -  
 DR InterPro: IPR002215; HLYD.  
 DR InterPro: IPR003997; RCLD.  
 DR Pfam: PF00529; HLYD; 1  
 DR PRINTS: PRO1490; RYTOXOIND  
 DR PROSITE: PS00543; HLYD.FAMILY; UNKNOWN\_1  
 SQ SEQUENCE 436 AA; 48443 MW; F9C4C68093D96563 CMC6;

Query Match	91.1%;	Score 1972.5;	DB 2;	Length 436;
Best Local Similarity	93.1%;	Pred. No. 6.7e-117;		
Matches 406; Conservative	3;	Mismatches 26;	Indels 1;	Gaps 1

[illegible]

## RESULT

ID	PRELIMINARY;	PRT;	436 AA.
Q92NX2			

DT 01-DEC

DT 01-DEC-2001 (Tremblé et al., 1999, Last annotation update)

GN SMC04208.

OC Bacteria; Proteobacteria; Bacteroidetes; Firmicutes; Actinobacteria; Cyanobacteria; Chloroflexi; Gemmatimonadetes; Planctomycetes; Thaumarchaeota; Euryarchaeota; Ciliophora; Metazoa; Fungi; Ascomycota; Basidiomycota; Mucoromycota; Zygomycota; Microsporidia; Viruses; Bacteriophages; Eukaryotic viruses; Archaeal viruses; Plant viruses; Fungal viruses; Retroviruses; Herpesviruses; Adenoviruses; Picornaviruses; Rotaviruses; Coronaviruses; Astroviruses; Bunyaviruses; Flaviviruses; Hepadnaviruses; Hepatitis B virus; HIV-1; HIV-2; HIV-3; HIV-4; HIV-5; HIV-6; HIV-7; HIV-8; HIV-9; HIV-10; HIV-11; HIV-12; HIV-13; HIV-14; HIV-15; HIV-16; HIV-17; HIV-18; HIV-19; HIV-20; HIV-21; HIV-22; HIV-23; HIV-24; HIV-25; HIV-26; HIV-27; HIV-28; HIV-29; HIV-30; HIV-31; HIV-32; HIV-33; HIV-34; HIV-35; HIV-36; HIV-37; HIV-38; HIV-39; HIV-40; HIV-41; HIV-42; HIV-43; HIV-44; HIV-45; HIV-46; HIV-47; HIV-48; HIV-49; HIV-50; HIV-51; HIV-52; HIV-53; HIV-54; HIV-55; HIV-56; HIV-57; HIV-58; HIV-59; HIV-60; HIV-61; HIV-62; HIV-63; HIV-64; HIV-65; HIV-66; HIV-67; HIV-68; HIV-69; HIV-70; HIV-71; HIV-72; HIV-73; HIV-74; HIV-75; HIV-76; HIV-77; HIV-78; HIV-79; HIV-80; HIV-81; HIV-82; HIV-83; HIV-84; HIV-85; HIV-86; HIV-87; HIV-88; HIV-89; HIV-90; HIV-91; HIV-92; HIV-93; HIV-94; HIV-95; HIV-96; HIV-97; HIV-98; HIV-99; HIV-100; HIV-101; HIV-102; HIV-103; HIV-104; HIV-105; HIV-106; HIV-107; HIV-108; HIV-109; HIV-110; HIV-111; HIV-112; HIV-113; HIV-114; HIV-115; HIV-116; HIV-117; HIV-118; HIV-119; HIV-120; HIV-121; HIV-122; HIV-123; HIV-124; HIV-125; HIV-126; HIV-127; HIV-128; HIV-129; HIV-130; HIV-131; HIV-132; HIV-133; HIV-134; HIV-135; HIV-136; HIV-137; HIV-138; HIV-139; HIV-140; HIV-141; HIV-142; HIV-143; HIV-144; HIV-145; HIV-146; HIV-147; HIV-148; HIV-149; HIV-150; HIV-151; HIV-152; HIV-153; HIV-154; HIV-155; HIV-156; HIV-157; HIV-158; HIV-159; HIV-160; HIV-161; HIV-162; HIV-163; HIV-164; HIV-165; HIV-166; HIV-167; HIV-168; HIV-169; HIV-170; HIV-171; HIV-172; HIV-173; HIV-174; HIV-175; HIV-176; HIV-177; HIV-178; HIV-179; HIV-180; HIV-181; HIV-182; HIV-183; HIV-184; HIV-185; HIV-186; HIV-187; HIV-188; HIV-189; HIV-190; HIV-191; HIV-192; HIV-193; HIV-194; HIV-195; HIV-196; HIV-197; HIV-198; HIV-199; HIV-200; HIV-201; HIV-202; HIV-203; HIV-204; HIV-205; HIV-206; HIV-207; HIV-208; HIV-209; HIV-210; HIV-211; HIV-212; HIV-213; HIV-214; HIV-215; HIV-216; HIV-217; HIV-218; HIV-219; HIV-220; HIV-221; HIV-222; HIV-223; HIV-224; HIV-225; HIV-226; HIV-227; HIV-228; HIV-229; HIV-230; HIV-231; HIV-232; HIV-233; HIV-234; HIV-235; HIV-236; HIV-237; HIV-238; HIV-239; HIV-240; HIV-241; HIV-242; HIV-243; HIV-244; HIV-245; HIV-246; HIV-247; HIV-248; HIV-249; HIV-250; HIV-251; HIV-252; HIV-253; HIV-254; HIV-255; HIV-256; HIV-257; HIV-258; HIV-259; HIV-260; HIV-261; HIV-262; HIV-263; HIV-264; HIV-265; HIV-266; HIV-267; HIV-268; HIV-269; HIV-270; HIV-271; HIV-272; HIV-273; HIV-274; HIV-275; HIV-276; HIV-277; HIV-278; HIV-279; HIV-280; HIV-281; HIV-282; HIV-283; HIV-284; HIV-285; HIV-286; HIV-287; HIV-288; HIV-289; HIV-290; HIV-291; HIV-292; HIV-293; HIV-294; HIV-295; HIV-296; HIV-297; HIV-298; HIV-299; HIV-300; HIV-301; HIV-302; HIV-303; HIV-304; HIV-305; HIV-306; HIV-307; HIV-308; HIV-309; HIV-310; HIV-311; HIV-312; HIV-313; HIV-314; HIV-315; HIV-316; HIV-317; HIV-318; HIV-319; HIV-320; HIV-321; HIV-322; HIV-323; HIV-324; HIV-325; HIV-326; HIV-327; HIV-328; HIV-329; HIV-330; HIV-331; HIV-332; HIV-333; HIV-334; HIV-335; HIV-336; HIV-337; HIV-338; HIV-339; HIV-340; HIV-341; HIV-342; HIV-343; HIV-344; HIV-345; HIV-346; HIV-347; HIV-348; HIV-349; HIV-350; HIV-351; HIV-352; HIV-353; HIV-354; HIV-355; HIV-356; HIV-357; HIV-358; HIV-359; HIV-360; HIV-361; HIV-362; HIV-363; HIV-364; HIV-365; HIV-366; HIV-367; HIV-368; HIV-369; HIV-370; HIV-371; HIV-372; HIV-373; HIV-374; HIV-375; HIV-376; HIV-377; HIV-378; HIV-379; HIV-380; HIV-381; HIV-382; HIV-383; HIV-384; HIV-385; HIV-386; HIV-387; HIV-388; HIV-389; HIV-390; HIV-391; HIV-392; HIV-393; HIV-394; HIV-395; HIV-396; HIV-397; HIV-398; HIV-399; HIV-400; HIV-401; HIV-402; HIV-403; HIV-404; HIV-405; HIV-406; HIV-407; HIV-408; HIV-409; HIV-410; HIV-411; HIV-412; HIV-413; HIV-414; HIV-415; HIV-416; HIV-417; HIV-418; HIV-419; HIV-420; HIV-421; HIV-422; HIV-423; HIV-424; HIV-425; HIV-426; HIV-427; HIV-428; HIV-429; HIV-430; HIV-431; HIV-432; HIV-433; HIV-434; HIV-435; HIV-436; HIV-437; HIV-438; HIV-439; HIV-440; HIV-441; HIV-442; HIV-443; HIV-444; HIV-445; HIV-446; HIV-447; HIV-448; HIV-449; HIV-450; HIV-451; HIV-452; HIV-453; HIV-454; HIV-455; HIV-456; HIV-457; HIV-458; HIV-459; HIV-460; HIV-461; HIV-462; HIV-463; HIV-464; HIV-465; HIV-466; HIV-467; HIV-468; HIV-469; HIV-470; HIV-471; HIV-472; HIV-473; HIV-474; HIV-475; HIV-476; HIV-477; HIV-478; HIV-479; HIV-480; HIV-481; HIV-482; HIV-483; HIV-484; HIV-485; HIV-486; HIV-487; HIV-488; HIV-489; HIV-490; HIV-491; HIV-492; HIV-493; HIV-494; HIV-495; HIV-496; HIV-497; HIV-498; HIV-499; HIV-500; HIV-501; HIV-502; HIV-503; HIV-504; HIV-505; HIV-506; HIV-507; HIV-508; HIV-509; HIV-510; HIV-511; HIV-512; HIV-513; HIV-514; HIV-515; HIV-516; HIV-517; HIV-518; HIV-519; HIV-520; HIV-521; HIV-522; HIV-523; HIV-524; HIV-525; HIV-526; HIV-527; HIV-528; HIV-529; HIV-530; HIV-531; HIV-532; HIV-533; HIV-534; HIV-535; HIV-536; HIV-537; HIV-538; HIV-539; HIV-540; HIV-541; HIV-542; HIV-543; HIV-544; HIV-545; HIV-546; HIV-547; HIV-548; HIV-549; HIV-550; HIV-551; HIV-552; HIV-553; HIV-554; HIV-555; HIV-556; HIV-557; HIV-558; HIV-559; HIV-560; HIV-561; HIV-562; HIV-563; HIV-564; HIV-565; HIV-566; HIV-567; HIV-568; HIV-569; HIV-570; HIV-571; HIV-572; HIV-573; HIV-574; HIV-575; HIV-576; HIV-577; HIV-578; HIV-579; HIV-580; HIV-581; HIV-582; HIV-583; HIV-584; HIV-585; HIV-586; HIV-587; HIV-588; HIV-589; HIV-590; HIV-591; HIV-592; HIV-593; HIV-594; HIV-595; HIV-596; HIV-597; HIV-598; HIV-599; HIV-600; HIV-601; HIV-602; HIV-603; HIV-604; HIV-605; HIV-606; HIV-607; HIV-608; HIV-609; HIV-610; HIV-611; HIV-612; HIV-613; HIV-614; HIV-615; HIV-616; HIV-617; HIV-618; HIV-619; HIV-620; HIV-621; HIV-622; HIV-623; HIV-624; HIV-625; HIV-626; HIV-627; HIV-628; HIV-629; HIV-630; HIV-631; HIV-632; HIV-633; HIV-634; HIV-635; HIV-636; HIV-637; HIV-638; HIV-639; HIV-640; HIV-641; HIV-642; HIV-643; HIV-644; HIV-645; HIV-646; HIV-647; HIV-648; HIV-649; HIV-650; HIV-651; HIV-652; HIV-653; HIV-654; HIV-655; HIV-656; HIV-657; HIV-658; HIV-659; HIV-660; HIV-661; HIV-662; HIV-663; HIV-664; HIV-665; HIV-666; HIV-667; HIV-668; HIV-669; HIV-670; HIV-671

NCBI TAXID=382;  
[1]  
OX  
PM

RE SEQUENCE FROM N.A.  
BC STRAIN=1021:

RA MEDLINE-213008234; EMBASE-11174104  
Galibert F.; Finan T.M.; Long S.R.

Boutry M., Bowser L., Cadieu E., Capela D., Chain P.

Gloux S., Godrie T., Goffeau A., Golding B., Guzy J., Gurjal M.,

RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,



RA Ramsperger U., Surzky R., Theault P., Vandebol M.  
RA Ramsperger U., Weidner S., Wells D. H., Wong K.-C., Batut J.  
RT "The composite genome of the legume symbiont *Sinorhizobium meliloti*."  
RL Science 293:666-672(2001)  
Q0 EMBL: AL591789, CMC4615.1; -.  
KW Complete proteome.  
Q0 SEQUENCE 436 AA; 47877 MW; 311E8A8EF3180C872 CRC64;

Query Match	28.0%;	Score 606;	DB 16;	Length 436;
Best Local Similarity	-34.0%;	Pred. No. 1.5e-30;		
Matches 145;	Conservative 75;	Mismatches 180;	Indels 26;	Gaps 6;

QY	24	IATFVLLG--MAAFAPLDSAVIANG--VYSAEYSDOVHLEGGMLAKIVRGEKKAG	80
Dd	23	VGLALISAAGMAKATTTQIAGAVYASGNEFYVDSHYK--VQHPITGCVIGELLIVEGDMYKAG	81
QY	81	QVLFELDPQANMAAGITRNQVYALKAMEARLLAERDQPSISFPADLTQSORDPWYARA	140
Dd	82	QVLMRLRDPQTKANLRIVTRKDELIMARLLAEERDDRGEEFFPELTKVAKERAAYASA	141
QY	141	IADQQAQFERRRITQGVNDLMNAQRLQYQSELEGIIDRQVQGLKDQIGFIEDELIDRLK	200
Dd	142	INGSKLTKEYKRNSREBRSSQLERRIVQIOLHETAGIDAOQAVYDRGISTILDAETASLRSL	201
QY	201	YDKGLVPRPLLALEARAGSLSGISGRITLDRSKAAQVAGADTOLKVPYKQEFFEYQS	260
Dd	202	HERGIVSVQNLNETLEPMTAFEGIRGKIKIAQQAQAGRIAEFTNLQIIDLDEDLKTEVRE	261
QY	261	ITETFRVLAVENTEKEVYASQAQRKIKIVSPVNGTQANLRFETBGAVYRAAPLYDIAPED	320
Dd	262	LREVOALGSEFTTERKIAAEQOLRIRIMLAQGGIVHOLNHTGVGISPADVYMSIYDPS	321
QY	321	EAFVIAQHPRPTQVDVNHMGMTVEVRLPAPHSAGNPRPERKHNDVAAVADRISDPQKARLF	380
Dd	322	DLRLATEKRIRPDIDQIQIOLQODVYLRSAQNRQTTREPLSGYSRIADLITEDPESGLSY	381
QY	381	LGIVRYDV-----KQRPNLHGRVYTAGMNPAAVIVPTGSEYRVLYQVLFSPRLTLR	429
Dd	382	--VYRVAVPAGAEALRLRGLSLEPLP-----GMPAAAFIQTQGRATLSTVLINPLSQIS	430
QY	430	TTMREE 435	
Dd	431	RAFRE 436	

RESULT 4	
09X6N7	
ID 09X6N7	PRELIMINARY; PRT; 436 AA.
AC 09X6N7;	
DT 01-NOV-1999 (TrEMBLrel. 12, Created)	
DT 01-NOV-1999 (TrEMBLrel. 12, last sequence update)	
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)	
DE RHIZOBIOICIN SECRETION PROTEIN RSPE.	
GN RSPE.	
OS Rhizobium leguminosarum (bivar trifolii).	
OG Plasmid pRie162Y10C.	
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group	
OC Rhizobiaceae; Rhizobium.	
OX NCBI_TaxID=386;	
ON [1]	
RN SEQUENCE FROM N.A.	
RP STRAIN=162Y10;	
RC MEDLINE=99318631; PubMed=10388672;	
RX Oresnik I.J., Twelker S., Oresnik M.F.;	
RA "Cloning and characterization of a Rhizobium leguminosarum gene	
RT encoding a bacteriocin with similarities to RTX toxins.";	
RL Appl. Environ. Microbiol. 65:2833-2840(1999).	
RN [2]	
RP SEQUENCE FROM N.A.	
RC STRAIN=162Y10.	
RX MEDLINE=21360352; PubMed=11467725;	
RA Venter A.P., Twelker S., Oresnik I.J., Hynes M.F.;	

RT "Analysis of the genetic region encoding a novel rhizobiocin from  
Rt *Rhizobium leguminosarum* bv. *viciae* strain 306.";  
Rh *Rhizobium leguminosarum* bv. *viciae* strain 306.";  
RL Can. J. Microbiol. 47:495-502(2001).  
DB EMBL: AF141032. A023601 1..

```
DR InterPro; IPR002215; HLYD.  
DR InterPro; IPR003997; RtxD.  
DR Pfam; PF00529; HLYD; 1.  
DR PRINTS; PR01490; RTXTOXIND  
DR Pfam; PF00529; HLYD; 1.
```

SQ SEQUENCE 436 AA; 47419 MW; 0CAFE93E641BAD75E CRC64;

Query Match	27.2%	Score 589;	DB 2;	Length 436;
Best Local Similarity	33.9%	Pred. NO. 1.7e-29;		
Matches 148;	Conservative 79;	Mismatches 198;	Indels 12;	Gaps 7;

QY	4	PKIOPRDNFCQAVNARIGGIIATLFEVGLIAPAPAPLDSAVIANG--VYSAEVSODVOHLE	62
Db	7	PSARAIRNLN-TVAALGTVLLGGVGM--GLAATTRLISGAVIASGTVVSDSYKPYVOHOK	63
QY	63	CGMLAKTLIVREGEKVVAGQVLFELDPQANAAGITENQYVALKAMBARLAEERDQPSI	122
Db	64	GGTGTGQIFVKKGDILVEAGQVLIHDDPTQTRANLAIYVRKLDELISARFARLAEERADGTV	123
QY	123	SEFPADLTISQARADPVYARINDEQAOFTERRQTLGGQVUDLMAQCLQYQSETEGIDRTQG	182
Db	124	EFSKQDTAKANDPOVSSITEEKKFLPADRRSSRMSKRAQLERLEQLOKQTEGVLVAETG	183
QY	183	LKDQLETEDELIDRLKYDGLVPRRRLLALEARAGSLSGSITRLADRSKAQOGASDT	242
Db	184	KRQSIGLIEKELSEJQRLPFGVLPAVRVYALQREANLNLGELGSLLANEQAKGRITET	243
QY	243	QLAKRQLOKEFFEVSOSITETPRRLAEVTEKEVVASDAQRIKIVSPVNGTQANLREFT	302
Db	244	ELQITQIDDDLRSEVSQULROAESDIEEFSERLVAEDDDLKRIDYVRSOGGVNHQLAVNA	303
QY	303	EGAVVRAAEPLVDIAPDEAEFVIOAHNQPDPVDVNHGMYTEVRLPFAHSAGNDPERHD	362
Db	304	PGAVIAGEALIMQVLPBDRLALAEVUKSPDIDQVITIGQVYHILFSAFSQSRST--PELNG	363
QY	363	PVA--VADRISSDPQKARLGLIVRVDKOLPRHLGRVY--AGMAPQIVTPGERTVQ	418
Db	362	VVASIADLVLAD--QSRGSLSYVVRAEVSDEGMORLQGVTPVPMQPVAEATIQGCEFTALA	419
QY	419	YLFSPLDLTLRTMREE	435
Db	420	YLLKPEMDVARRAFKEE	436

RESULT	5
033679	
ID	033679
AC	033679; PRELIMINARY; PRT; 439 AA.
DT	01-JAN-1998 (TREMBLrel_05, Created)
DT	01-JAN-1998 (TREMBLrel_05, Last sequence update)
DT	01-DEC-2001 (TREMBLrel_19, Last annotation update)
DE	PRSE (PROTEIN SECRETION PROTEIN, HLYD FAMILY).
OS	<i>Rhizobium meliloti</i> (SinoRhizobium meliloti).
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiales group;
OC	Rhizobiaceae; SinoRhizobium.
OX	NCBI_TaxID=382;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Tork G.M.; Walker G.C.;
RT	"The <i>Rhizobium meliloti</i> exoX gene and prsD/prsE/exsH genes encode
RT	components of independent degradative pathways which contribute to
RT	production of low-molecular-weight succinoglyca.";
RL	mol. Microbiol. 25:117-134(1997).
RP	SEQUENCE FROM N.A.
RC	STRAIN=1021; PLASMID=PSYMB (MEGAPLASMID 2);
XX	MEDLINE=21396508; PubMed=11481431;

RA Flihar T.M., Weidner S., Wong K., Buhrmester J., Chan P.,  
RA Vonholster F.J., Hernandez-Lucas I., Becker A., Gouzy J.,  
RA Golding B., Puelhner A.:  
RT "The complete sequence of the 1,683-kb pSMB megaplasmid from the N2  
RT fixing endosymbiont *Sinorhizobium meliloti*,"  
RT Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).  
RL EMBL; U89163; AAB64092.1; "-."  
DR EMBL; AL603646; CAC94691.1; "-."  
DR InterPro: IPR002215; HLYD.  
DR InterPro: IPR003997; REXD.  
DR Pfam: PF00529; HLYD; 1  
DR PRINTS; PR01490; RYTXOXIND.  
DR Plasmid; Complete proteome  
KW SEQUENCE: 439 AA; 48036 MW; 82B2421825522BBD CRC64;

Query Match	26.9%	Score 583	DB 16	Length 439
Best Local Similarity	31.9%	Pred. NO. 4.2e-29		
Matches 134; Conservative	84;	Mismatches 198;	Indels 4;	Gaps 4

0y 19 IGYGIIATFEV-GLIIGMAFPIBDSAVIANVSAEVS-ODVOHLEGGMIAKLIVGEK 76

0y 21 IGVSVLALALVAVGGMAATTELISAIYAGSVIYVDNNKKYQHLTGIGYIGELLVNGDR 80

0y 77 VKAGOVFEELDPQANAAAGITRNOYALAKAMEARLLAERDQRESIFPADLTSQRADP 136

Db 81 VBAGOVLLRDOTTVRAMALIIESTIAQFVARRARLOAERMGASFEIEDIAEPIGTA 140

0y 137 VARAIADBOAQFETTEROITIOGVNDLMNAORLOQYSEIEGIDROTGAKDOLGFIEDBLID 196

Db 141 AAKLIEGEORLEFASRRSALSCKMGOLDSRKQOLADEVEGLVQJNAIEALKLIAEELTG 200

0y 197 LAKLVKGIVPRRLALAEARASISGSGIGLTLADRSKAYGASDPTOLKVRIOKEEFFQ 256

Db 201 VOSLEGGIVPEORVYTLTKROPAELEGGRRHHIARRAOKSSSEIDLIQIOLDEDRSE 260

0y 257 VQSQITETFEVRLAEYTEKEVVASDAOKRIKIVSPNQTACQIMRFETEGAVVRAAEPLVDI 316

Db 261 ISKELTIDVAKIAEVEBERTATATDOLRLDITAPLSGIIYOLAIHTYNGVINPEETIMLV 320

0y 317 ABEDAEVIOAHFOPTDVDDNVMHGVTEVRLPAFHSAGNDPBERHDPVAAADISDPQO 376

Db 321 VBEAEDLVEAKVATAHNDIOLRVGQSVIEIRSFANQRTPEVEAEVUYVADPLTDBERTG 380

0y 377 ARFLGIVAVNDYKOLPRLRG-RVYANMPAOYIYVGERTLQVLESPRLBTLTMBEE 433

Db 381 ASYVPLRIPPKAESI-AKTKGSLIPGMAEVEFKIARVYIYSTVRLTDIOMHAFRED 439

RESULT	6
09HYJ9	
ID	09HYJ9
AC	09HYJ9;
DT	01-MAR-2001 (TREMBLrel, 16, Created)
DT	01-MAR-2001 (TREMBLrel, 16, last sequence update)
DT	01-OCT-2001 (TREMBLrel, 18, last annotation update)
DE	METALLOPROTEASE SECRETION PROTEIN.
GN	HASE OR PA3405.
OS	Pseudomonas aeruginosa.
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC	Pseudomonas.
OX	NCBI_TaxID=287;
RN	{1}
RP	SEQUENCE FROM N.A.
RC	STRAIN=ATCC 15692 / PA01;
RX	MEDLINE=20437337; PubMed=10984043;
RA	Stover C.K., Plam X.-O.T., Erwin A.L., Mizoguchi S.D., Warren
RA	Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Jaseg
RA	Barber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuen
RA	Grady L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA	Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT	Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen." ;  
 RL Nature 406:959-964 (2000) .  
 DR EMBL: AE004761: AAC60793.1: -.  
 DR InterPro: IPR002315: HLVD.  
 DR InterPro: IPR003997: RCD.  
 DR Pfam, PF00529: HLVD: 1  
 DR PRINTS, PR01490: RYTXOXND.  
 DR Protease; Complete Proteome  
 KW SEQUENCE 443 AA: 48776 MW: 06CCD3E9DE7092A7 CRC64

Query Match	26.1%;	Score 564;	DB 16;	Length 443;
Best Local Similarity	32.5%;	Pred. No. 6.7e-28;		
Matches 138;	Conservative 87;	Mismatches 189;	Indels 10;	Gaps 4;

```

0Y 17 ARIGGIIATLFEVGLGMAAFAPIDSANVINGVYSAEVSOD-VONHEGGMLAKILYREGG 75
Db 25 SRLGGIVLLFVGFLLMAGLAPLDKGVSGVIVMAGSKKAYOHHPGGIVRIIRVHEGE 84
0Y 76 KYKAGVQVFEELDPQANMAAGITRNOYVAKAMEARLLAREBDRSPISFPADLTOSRAPD 135
Db 85 RVEAGQVFLLEMDATQARQADGFLPQYUALLALASLARLSAEDEKARIEPAELLA-LDDP 143
0Y 136 MYARAIADGQAFTERRTOTGGQYVDLMAAQRLQYQSEIEGIDRQOTGLKQGLFIDEELI 195
Db 144 RLPTLLEOOROLHDSRRRLRLLELDGIAETVAGSQAOLDGQAALRSKEORPALEBOLR 203
0Y 196 DLRLKYDGLGPRPLALLAERA----GSLGSGIRLTADRSKAAQVQASDTOLKVRIOKO 251
Db 204 GLRQLASGCVPRNRLLDSEKRLLAQVNEINGDLGSLGSTRQOLL-----ELRLRMQORRE 259
0Y 252 EFFEQVQSITETVRVLAETVEKEVYASDAOKRIKIVSPVNGTAKONIREFTEGAVYRAAE 311
Db 260 KFOEYRASLADQVRAEELRNRLASAFEDLANSEVVRPAAGLVGQEVTEGCVLAPGQ 319
0Y 312 PLVYDAPDEAFVIOAHQOPFDVDVHVHGCVTEVRLPAPFHSAGNPDEERHDPAVADARIS 371
Db 320 QLMETLPEKQPLVYARLPKEMVKDKVYRGLVVELMFSAGNSSTTPRVEGVTLYVSADRLL 379
0Y 372 DPQOKARFLGIVDVVQOLPRPHLKGRTVAGMPAQVIVPTGERTVLOLYLESPLDPLRTT 431
Db 380 DERSEAPYRVIRIVRGEGBVRRLAGLITRCPMPVEAFVRSSEERSILNYLFKRLPADRTHLA 439
0Y 432 MREE 435
Db 440 LGBE 443

```

```

RESULT 7
ID 098LC9 PRELIMINARY; PRT; 434 AA.
AC 098LC9;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE RHIZOBIOTICIN SECRETION PROTEIN, RBEPE.
GN ML1026.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).

```



DR InterPro: IPR0023215; HLYD.  
DR InterPro: IPR003997; RtxD.  
DR Pfam: PF00529; HLYD\_1  
DR PRINTS: PR01490; RLYTOXIND.  
DR PROSITE: PS00543; HLYD\_FAMILY; 1.  
KW Protease.  
SEQUENCE 443 AA; 48409 MW; 762FED7F822BB0C44 CRC64;

Query Match	25.08	Score 540.5	DB 2	Length 443
Best Local Similarity	31.78	Pred. No. 2.1e-26		
Matches 138; Conservative	85	Mismatches 178	Indels 35	Gaps 6

QY	18	RIGIGIALTIVGILGMAAFPLDUSAV-IANGVISAESODVCHLEGCMATKILYREGK	76
Db	25	RIGGVLVLLGGGGLGMLPLPDKGVPVSGSVVVAAGNRRKAVOHPSGGVSQLOVHEGDR	84
QY	77	VKAGOVLEELPPTONAAAGITRRNOYALKMEARLLAEORDRESIFFPADLTQSADPM	136
Db	85	VKAGOVLLIMDTVYSRQRODLRSQOGLSMAOQARLOAEROGHAIYFPALLQARREPE	144
QY	137	VARAIDEQAOFTERROTITOGVDLMMNROLYOSEIEGIDROTGKLDQGFIEDELID	196
Db	145	VMSLMLLOOQOFTSRRAILOSELAIRESISGOSAMEGIESQSVASROKRAMIOEOIGG	204
QY	197	LRIKYDKGLVPRPRLALE-----ARAGSLSGISGIRLADRSKANQOASDPOKVKROI	249
Db	205	MKRLAAGSVVARNRLLDLDEGQYAOIDGQASDSDTNIGRL-----GRQILELTKIRAI	255
QY	250	--KOEFFBVSQSITETFRVLRALEVTEKEVVASDAOKRIKIVPYNGTAQMLREFTEGAVV	307
Db	256	ORREYQKEVSSQJLAEVBMKLDDELINRLKAEADLGHQYKASVAGTVGSLVTEGCVI	315
QY	308	RAAEVLVUIADEDEFVIGAHFOPTVDVNVMGAVTEVRLPAHFSAGNPPPERIKDPAVA	367
Db	316	GAGQOLMIVTSDRGLOVEARIYVELLIDKOVAGPLVELLSAFMPSSTTPRVEGEVTLVGA	375
QY	368	DRISDPQKQARLF-----LGIVRVVDKOLPPLHNGRVATGAPPAOIVVPTGERTVLOY	419
Db	376	DRLTDEKSGAPYYSVRKAVSEAGIQRNLGILIRP-----GNPVGSGFIRTGSRSMNY	427
QY	420	LFSPRLDPLRTTMRKE	435
Db	428	LEKPLTDRLHALTEE	443

```

RESULT 10
005198
ID 005198 PRELIMINARY; PRT; 473 AA.
AC 005198;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MEMBRANE FUSION PROTEIN.
GN EXPD2.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EFBI;
RC MEDLINE=98162063; PubMed=9501442;
RA Lloret J., Wulff B.B.H., Rubio J.M., Downie J.A., Bonilla I.,
RA Rivilla R.;
RT "Exopolysaccharide II production is regulated by salt in the
RT halotolerant strain Rhizobium meliloti EFBI."
RL Appl. Environ. Microbiol. 64:1024-1028(1998).
DR EMBL: Y08703; CAB41456.1;
DR InterPro: IPR002215; HlyD.
DR InterPro: IPR003997; RcxD.
DR Pfam: PF00529; HlyD; 1.
DR PRINTS: PR01490; RTYTOXIND.

```

SQ SEQUENCE 473 AA; 51603 MW; D536469E385319FF CRC64;

Query Match:	24.8%:	Score 536.5:	DB 2:	Length 473:
Best Local Similarity	28.5%:	Pred. No. 4e-26:		
Matches 119, Conservative	96:	Mismatches 199:	Indels 3:	Gaps 2

QY	20	GXGIIALFEVCLLMAAFPLDSVINGVVA5V5--ODVUHEGGGLAKIVREGKXV	78
Db	59	GLTITLVAFGFPCGMASTELSSASVSGTIVDSKRTYSHEGCVSLRVQBGDVA	118
QY	79	AGOVLFELDPQANAAGITFNOYVAKAMEBARLAEORDRPSISFPADLTISQADPMA	138
Db	119	PGQPLMQJEDTRANSDDQALESSRRVGLIAKLARLSIELAGIQAVIDFPDDLVA--AGAA	176
QY	139	RAIADQAOFTERRTOTGOVDLMAORLOYSIEGIDROTQGLQOLFIEDELIDLR	198
Db	177	DAVTAETAFEEKRNDEAGRTAIOKRTIEEYSEKAKSLTQLOATDROIELMBEORTAIA	236
QY	199	KLYOGVAPRRRLALEARASLGSIGRLTADSKKAVOGASDPTOLKVRIOKOFFEVS	258
Db	237	TLVEAFAPORSKRLAIDRUSLATAKBEIAGDAQAQAKAGAEHLTGISPOSEIA	296
QY	259	QSITETRYRLAEVIEKEVVASDAOKRTIKYSPVNGTQONLRFEEGAVVAAREPLVDIA	318
Db	297	GEITTARJELAEVERIRIISADVLRLEIRAPQGIANYATOLRPPGSAAVYPCQPLDI	356
QY	319	EDEAFVIAHFQPTDVDVNHMGVTEYVRLPAFHSAGNPDDERHDPVAADRIDSPOKOAR	378
Db	357	EDEPLLVEMHVSSTRDIDISITIGSSTOIRLTAVNORSHLPMDGKVTYIAADQSMDEK	416
QY	379	LFLGLVIRVADQORPPLHGRGRTAAMPQOVIYVPIGGERVLYOLYSPPLDLTIRTMREE	435
Db	417	YFVARAEIAPSLANDPIDIRLPCMPAEVILVHKARSAIDYLVSPSDSSNRAERED	473

```

RESULT      11
0921J2
ID      0921J2      PRELIMINARY;      PRT;      511 AA.
AC      0921J2;
DT      01-DEC-2001 (TREMBLrel. 19, Created)
DT      01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE      ALKALINE PROTEASE SECRETION PROTEIN ABRE.
GN      APRE OR RC0428.
OS      Rickettsia conorii.
OC      Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC      Rickettsiaceae; Rickettsiae; Rickettsia.
OX      NCBI_TaxID=781;
RN      [1]
RN      RP      SEQUENCE FROM N.A.
RN      RC      STRAIN-MALISH 7;
RN      RX      MEDLINE=21442074; Pubmed=1157893;
RA      Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA      Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA      Raoult D.;
RT      "Mechanisms of evolution in Rickettsia conorii and R. prowazekii."
RL      Science 293:2093-2098(2001) .
DR      EMBL; AE008606; AAU0266.1; -.
KW      Protease; Complete proteome.
SQ      SEQUENCE      511 AA; 57582 MW;      4BD9B8A55866A5C CRC64;

```

```

Query Match          24.5%; Score 530.5; DB 16; Length 511;
Best Local Similarity 30.0%; Pred. No. 11e-25;
Matches 135; Conservative 88; Mismatches 202; Indels 25; Gaps 7

0y      5 KIQRPMDNQAVAR--IGYIATLFPVGLG--WAAFAPLDSAVINGVAASEVS-QDVO 59
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      68 KIDKDRNNVAQARSTILFGIYVIMFVLVIGGLMSLAPLDSGAVAVGIVMPSINKKTD 127
      : : : : : : : : : : : : : : : : : : : : : : : :

0y      60 HLEGGLAKLIVREGKAVAGVLELFDPTQANAAAGITRNQVYALAMEARILLAEIRDOR 119
      : : : : : : : : : : : : : : : : : : : : : : : :

```

[illegible]

OY	128	HHEGGIIINAIYKQDCKKEGKLLEETFRKSHEHNIGQRYRFLATENKLIERNLT	187
OY	120	PSISFPADLTSORADPMVVARALADEQAQFTERRQTIGQVDLMNAQRLOYOSEIEGIDRQ	179
Dd	188	EQIEESDFELMODINLPVAKIITHGTENLFRSRKVEYNSEKDALHONIQLEKKIEGLEAK	247
OY	160	TQGLKDQLGFIEDLEIDLRLKLYDKLVPRPLLAEARAGSISGSLGRILTDRSAVAOGA	239
Dd	248	KVAASKTAEVYODRIKALRTLTKEGFVKRAALLDOEAKVAAKSRSVATTTEATIGIRHAI	307
OY	240	SDTOLKVQNIQOEPEFOYSOSTEFERVRVLAETEKEVVASDAQKRITKIVSPNGTANLR	299
Dd	308	TETQKIINOQKKITERLTLELRQAQVASTCEKYEMALTDSLNRYILRAPVDGYVNSLK	367
OY	300	FETEGAVVRAAEPLVDIAPEDEAFVIOAHFOPTDDNVHMGMATEVRLPAFSAGNP---	356
Dd	368	YHTIGCVISHQGPIAMEISPTNDPDLIEAKVSQKNIDSHEGIVAAIRSAFSSRTTPFT	427
OY	357	-DPENHDVVAAADRISDP-QKQARLFLCIVRD-----VKQLPHLGRVTAGMPA	405
Dd	428	GKVVISISDIYODEHQYEGQOQDNYYVARVELIDMEFNKVAKVNKLIEH-----PGMQA	481
OY	406	OVIPTGERTVLOYLFSPLRDLRTTMEB	435
Dd	482	EQVITGTIRLTRLLDPPVTDFAEKAREK	511
 RESULT 12 O9XB63 PRELIMINARY; PRI: 443 AA.			
ID	O9XB63	PRELIMINARY;	PRI: 443 AA.
AC	O9XB63		
DT	01-NOV-1999	(TREMBLrel. 12, Created)	
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DE		PRTE PROTEIN.	
GN		PRTE.	
OS		Erwinia amylovora.	
OC		Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;	
OC		Erwinia.	
OX	NCR1_Taxid=552;		
RN	[1]		
RP		SEQUENCE FROM N.A.	
RC		STRAIN-E9;	
RX	MEDLINE=99303693; PubMed=10373365;		
RA	Zhang Y., Bak D.D., Heid H., Geider K.;		
RT	"Molecular characterization of a protease secreted by Erwinia		
RT	amylovora.";		
RL	J. Mol. Biol. 289:1239-1251(1999).		
DR	EMBL; Y19002; CAB42875.1; -		
DR	InterPro: IPR002215; HLyd.		
DR	InterPro: IPR000847; HTM_LysR.		
DR	InterPro: IPR003997; RtxD.		
DR	Pfam: PF00529; HLyd.1.		
DR	PRINTS: PR01490; RTXTOKIND.		
DR	PROSITE: PS00543; HLyd_FAMILY.1.		
DR	PROSITE: PS00044; HTM_LYSR_FAMILY; UNKNOWN.1.		
SQ	SEQUENCE 443 AA; 48865 MM; 59AD285715994011 CRC64;		
 Query Match 24.1%; Score 522; DB 2; Length 443; Best Local Similarity 29.8%; Pred. No. 3e-25; Matches 137; Conservative 81; Mismatches 181; Indels 60; Gaps 9;			
OY	7	QRPNTNQAVARIIGCIATLTFVGILLGMAAFPLDSAVIANGVSAE-VSDQVHLLEGM	65
Dd	15	QIPQDE-RRYTTMGWLAGVGLEGLAMAAEPALDKGAASPTVYSGNSKTVQAPASI	73
OY	66	LAKIIVREGEKKYKAQOVLELDPTQANAAGACTRQAYALKAMEARLLAEORDORSISFP	125
Dd	74	IINIIVAKESGDKKAKAILVOLLSOVAQAOVDSLRRYITTLATIEGRLLAEARGMLNYVTS	133
OY	126	ADLTQSARDPVVARIADEQAQFTERRQTIGQVDLMNAQRLOYOSEIEGIDRQTGLEKD	185

Db	134	PFLA0VKQ0PRVVEIIAQTQLFASRROGLSEID-----GYQSDMGRNFOLKGLD	186
QY	186	-----QUGFIEDLIDLRKLYDGLVPRPRLALEARAGSL-----GSIGRLTAD	231
Db	187	SRANKQIOLSSIREQMSMKQLAADGYLPRNNRYLDVQGFQFAVSSSNETAGRIGQL---	243
QY	232	RSKAVQASDITOLKVRQIKQEFFEVSOSITETRYRLAEVTEKEVVASDAQRITVSPV	291
Db	244	-OKOLO---ESQORIDORFADYQREVRQTOLAQOTQMDISEFRNKLQMAPFDLGNATITSPV	299
QY	292	NGTAQNLRFRETCGAVVRAAEPLVDIADPEADFAVIAOHPTDYDVVHHGMTEVRLPAP	351
Db	300	DETIVGLNLFITQGVGAGDHLMDVPSOALVVDSHLKVLDLIDVYVQGLPYDLMTAFN	359
QY	352	SAGNDPDERHDPVAADRIDSDPOKQARLFLGIVRDVQOLPRLHGRVYA-----	401
Db	360	QNKTKRIGTGYLVASADR-----YDKNSGEPTYYMQVTVSPREGNMLSG	404
QY	402	-----GMPAQIVTPPGERTVQLYFLSPRLDRLTTRREE	435
Db	405	EDIKPMPEVEFVKTSRSLSTYLEKPLIDRAHTSLTEE	443
RESULT	13		
Q54457		PRELIMINARY; PRT: 443 AA.	
AC	Q54457:		
DT	01-NOV-1996 (TREMBLrel. 01. Created)		
DT	01-NOV-1996 (TREMBLrel. 01. Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19. Last annotation update)		
DE	LIPIc.		
GN	LIPIc.		
OS	Serratia marcescens.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
CC	Serratia.		
OX	NCBI_TaxID=615;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-SR41;		
RX	MEDLINE=96062219; PubMed=7592412;		
RA	Atsuka H., Kawai E., Omori K., Shibatani T.;		
RT	"The three genes lipB, lipC and lipD, involved in the extracellular		
RT	secretion of the Serratia marcescens lipase which lacks an N-terminal		
RT	signal peptide.";		
RL	J. Bacteriol. 177:6381-6389(1995).		
DR	EMBL: D49826; BAA08632.1; -.		
DR	InterPro: IPR002215; HLYD.		
DR	InterPro: IPR003997; RtxD.		
DR	Pfam: PF00529; HLYD: 1.		
DR	PRINTS: PR01490; RXTXOIND.		
DR	PROSITE: PS00543; HLYD_FAMILY: 1.		
SEQ	SEQUENCE 443 AA: 49000 MM; B69DEC70F474EB64 CXC64;		
Query Match	24.1%;	Score 521.5;	DB 2; Length 443;
Best Local Similarity	29.3%;	Pred. NO. 3.3e-25;	
Matches 133;	Conservative 88;	Mismatches 194;	Indels 39; Gaps
QY	6	IORPPTDNF-----QAVARIGYGIITLVGLLGMAAFAPLDSAVIANGVYSAEVS-Q	56
Db	5	IIEPDSDYEEELIPDERRTFRKMGWLVGGIFGLMAFAFLDGVASPGSVTSQGNK	64
QY	57	DYQHLGGGLAKIIVREGKVKAGOVLPFLDPTQANAAAGITRNOYVVALKAMEARLAE	116
Db	65	TVQAPASGIIKNIARADGDKVAGVELVQVSOVQAOQVDSLRDQYTTTLATEGRLLAE	124
QY	117	DORPSIPADLTISORADPMVARAIADAEQAFTEERRQITQGVDMNMQRLQOSEIEGI	176
Db	123	DELSTVTFSPILDVAKDKRVAEIIALQQLFASRRQAQSDID-----GYKQSMGI	177
QY	177	DRQTOGLKD-----QUGFIEDLIDLRKLYDGLVPRPRLALEARAGSLSGSIGRLT	229
Db	178	FRQLQGLDSDRGNKOIOLSSIREQMSMKQLAADGYLPRNNRYLDVQGFQFAVSSSNET	237

OY	230	ADRSKAVGASDITDLVKRQIKOEFEEOVSOSITETRVLAETEKEVVASDAQRIKIVS	289
Dd	238	GRIGQLQKQLLESQORIDQRPADYQREKRITQIAOTQMDSSEFRKLNOADDLGNATITTS	297
OY	290	PVNGTAQNILREFTEGAVVRAAEPLVDIAPDEFAEVIQAHPQPTDVNDVMGVTEVLPA	349
Dd	298	PVDGTIVGLNFTGGVVGAGDDHMDVPSQATLIVDSRLKVDELDKVYNGLPDLMPETA	357
OY	350	FHSAGNPPEKHDPVAADRISD-----PKQARLFLE---GIVRDVKQLPRPHLRGRVTA	401
Dd	358	FNOKTPKIPCTVTLVSDRLVLDKANEPYQOMOVTVSPSGMKMLSGEDIKP-----	409
OY	402	GMPAOVIVPTGERFVLYQIFSLPRDLTTTMRRE	435
Dd	410	GMPVEVFYKTGSRSLSLTLPFKPIIDRATITSLEE	443
RESULT	14		
O9ZDL4			
ID	09ZDL4	PRELIMINARY;	PRT; 511 AA.
AC	09ZDL4		
DT	01-MAY-1999	(TREMBLrel. 10. Created)	
DT	01-MAY-1999	(TREMBLrel. 10. Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19. last annotation update)	
DE	ALKALINE PROTEASE SECRETION PROTEIN APRE (APRE).		
GN	Rp314.		
OS	Rickettsia prowazekii.		
OC	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;		
CC	Rickettsiaceae; Rickettsiae; Rickettsia.		
OX	NCBI_taxid=782;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MADRID E:		
RX	MEDLINE=99039499; PubMed=9823893;		
RA	Anderesson S.G.E., Zomorodipour A., Andersson J.O.,		
RA	Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,		
RA	Eriksson A.-S., Winkler H.H., Kurland C.G.;		
RT	"The genome sequence of Rickettsia prowazekii and the origin of		
RT	mitochondria.";		
RL	Nature 366:133-140(1998).		
DR	EMBL, AJ235271; CAA14774.1; .-		
DR	HSSP; P10802; 11YU.		
DR	InterPro; IPR002215; HLyd.		
DR	InterPro; IPR003997; RKXD.		
DR	Pfam; PF00529; HLyd. 1.		
DR	PRINTS; PR01490; RTXTOXIND.		
DR	PROSITE; PS00543; HLyd_FAMILY; UNKNOWN_1.		
KW	Complete proteome.		
SO	SEQUENCE 511 AA; 57770 MW; CFC6EC6005BA34A8 CRC64;		
Query Match	24.0%;	Score 519.5;	DB 16; Length 511;
Best Local Similarity	29.3%;	Pred. No. 5.3e-25;	
Matches 134;	Conservative 93;	Mismatches 192;	Indels 39; Gaps 10.
OY	6	IORPTD---NFQAVAR--IGYGIALLTFVGLLG--MAAFAPLDSAVIANGVSAEVS-Q	56
Dd	65	IRKKTDENNNNAQAARSPIILFGIVIIFFVLIGGLMSALAPLDGGAVALGIVISTNK	124
OY	57	DVOHLEGGLAKILYREGCKVYAGOVLEFLDPTQANAAAQITRNQYVALKAMEARLLER	116
Dd	125	TIOHNHGCIINAIVYKOGDKVEGKGLTELLETIRIKSEHENITLSQYRNFLATENRLLAR	184
OY	117	DQRPSISFPADILTSQAD-PWVARIALDAQOAFTERROTIQGOVDMNAQRLOYOSEIG	175
Dd	185	DNLQEQKF-SDFLMONINLPEVAKIIHTDENLFSKSKEYESEKDAPNONIAOLEKKIDG	243
OY	176	IDROTGLKLDQGLGETDELIDLKLYDKGLVPRLPLALEARAGSLSGSIGRTLDRSKA	235
Dd	244	LEAKRIAAKSTSEVVODRLKALRTLKEGFIQKALLDEEAVVAAKSVATTAEAIAGI	303
OY	236	VQGASDTOLKVRHQIKOEFEEOVSOSITETRVLAETEKEVVASDAQRIKIVSPVNGTA	295

Db	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
Db	304 RHATTEVOIKLINQNKQTEBTETLIELRPAQIQOTASIKKYNLSLDDLNRVILIRSPVOGIV	23.6%; Score 512; DB 2; Length 435;	132;	71;	208;	18;	4;
Qy	296 QNLRFEAGAVVRAEPLVDIAPEDEAFVIOAHFOPTDVNVHMGWTEVRLPAFHSAGN	30.8%; Pred. No. 1.3e-24;	132;	71;	208;	18;	4;
Db	364 NNLKHTIGVISHQOPIEISPIINDPLIIEARISQKIDISVEGLVAKIFSAFKSKRT		132;	71;	208;	18;	4;
Qy	356 P-----DPERHDPVAVADRIDSPQKARFLGIVRYV-----VKOLRPHLHG		132;	71;	208;	18;	4;
Db	424 PFTGKVVSPD---IVQDEROHLGOQDQNVYARLEIMHEERNKAVKKNLALH---		132;	71;	208;	18;	4;
Qy	398 RVTAGMPAOVIVPTGERTVLOLIESPLDRLTTRREE		132;	71;	208;	18;	4;
Db	477 ---PQMAEVOIVTQTRTLRLTLRLDLPVDTAFKFAFKREK		132;	71;	208;	18;	4;
RESULT	15.						
ID	005694	PRELIMINARY;	PRT;	435 AA.			
AC	005694;						
DT	01-JUL-1997 (TREMBLrel. 04, Created)						
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)						
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)						
DE	PRESE PROTEIN.						
GN	PRESE.						
OS	Rhizobium leguminosarum.						
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;						
OC	Rhizobiaceae; Rhizobium.						
OX	NCBI_taxid=384;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=8401;						
RA	Flonie C., Hartley N.M., Findlay K.C., Downie J.A.;						
RT	"The Rhizobium leguminosarum prsB genes are required for secretion of						
RT	several proteins, some of which influence nodulation, symbiotic						
RT	nitrogen fixation and exopolysaccharide modification.";						
RL	Mol. Microbiol. 0:0-0(0).						
DR	EMBL; Y12758; CA73293.1;						
DR	InterPro; IPR002215; HLYD.						
DR	InterPro; IPR003997; REXD.						
DR	Pfam; PF00529; HLYD; 1.						
DR	PRINTS; PR01490; RTYTOXIND.						
SO	SEQUENCE 435 AA; 47533 MW; DA1F71655D96D131 CRC64;						

Mon Jul '22 09:14:02 2002

us-09-913-414-5.rspt

Page 9

[illegible]

Search completed: July 19, 2002, 10:48:43  
Job time: 418 sec

THIS PAGE BLANK (USPTO)